

**MANAGEMENT INFORMATION EXTRACTING DEVICE EXTRACTING
MANAGEMENT INFORMATION FROM IMAGE AND ITS METHOD**

Patent Number: JP10240958
Publication date: 1998-09-11
Inventor(s): NAOI SATOSHI; KATSUYAMA YUTAKA; TAKEBE HIROAKI
Applicant(s): FUJITSU LTD
Requested Patent: ☐ JP10240958
Application Number: JP19970346849 19971216
Priority Number(s):
IPC Classification: G06T11/60; G06F19/00; G06F17/30; G06K9/20
EC Classification:
Equivalents:

Abstract

PROBLEM TO BE SOLVED: To extract management information proper for managing an image from the optional inputted image.

SOLUTION: A management information extracting device learns the ruled line structure of a document A and a title position designated by a user at the time of learning a form and houses them in a layout dictionary 31. Then, at the time of operating, the device extracts the ruled line structure from the image in an unknown document 32 to match with a document within the dictionary 31. When the document 32 and the document A coincide with each other, e.g. information on a title position designated by the document A is referred to extract a title from a character string area 33 in the document 32.

Data supplied from the esp@cenet database - I2

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平10-240958

(43) 公開日 平成10年(1998) 9月11日

(51) Int. Cl. ⁶
 G06T 11/60
 G06F 19/00
 17/30
 G06K 9/20 340

F I
 G06F 15/62 325 D
 G06K 9/20 340 C
 G06F 15/22 C
 15/401 310 A

審査請求 未請求 請求項の数42 O L (全36頁)

(21) 出願番号 特願平9-346849

(22) 出願日 平成9年(1997)12月16日

(31) 優先権主張番号 特願平8-351346

(32) 優先日 平8(1996)12月27日

(33) 優先権主張国 日本 (J P)

(71) 出願人 000005223

富士通株式会社

神奈川県川崎市中原区上小田中4丁目1番
1号

(72) 発明者 直井 聡

神奈川県川崎市中原区上小田中4丁目1番
1号 富士通株式会社内

(72) 発明者 勝山 裕

神奈川県川崎市中原区上小田中4丁目1番
1号 富士通株式会社内

(72) 発明者 武部 浩明

神奈川県川崎市中原区上小田中4丁目1番
1号 富士通株式会社内

(74) 代理人 弁理士 大菅 義之 (外1名)

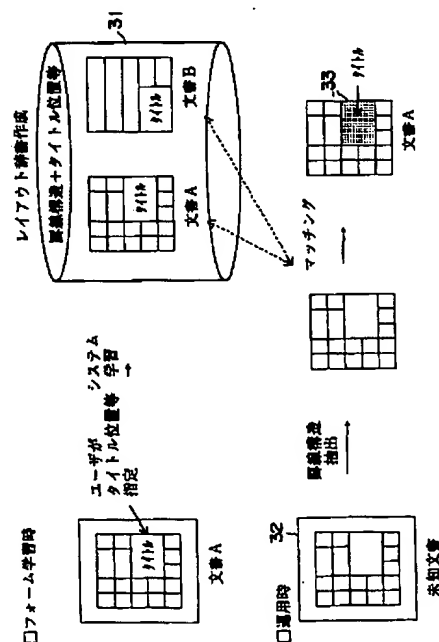
(54) 【発明の名称】 画像から管理情報を抽出する管理情報抽出装置および方法

(57) 【要約】

【課題】 任意の入力画像から、その画像を管理するための適切な管理情報を抽出することが課題である。

【解決手段】 管理情報抽出装置は、フォーム学習時に、文書Aの罫線構造とユーザが指定したタイトル位置を学習し、レイアウト辞書31に格納する。そして、運用時に、未知文書32の画像から罫線構造を抽出し、レイアウト辞書31内の文書とのマッチングを行う。例えば、文書32と文書Aが一致した場合、文書Aに指定されたタイトル位置の情報を参照して、文書32の文字列領域33からタイトルを抽出する。

管理情報抽出処理を示す図



【特許請求の範囲】

【請求項 1】 任意の入力画像に含まれる表領域の外郭部分に対する罫線の相対的な位置情報に基づいて、該入力画像に含まれる管理情報の位置を算出する計算手段と、

前記計算手段が算出した位置に基づいて、前記入力画像から前記管理情報を抽出する抽出手段とを備えることを特徴とする管理情報抽出装置。

【請求項 2】 前記計算手段は、前記表領域の外郭部分の情報として、該表領域の基準となる大きさと該表領域の外郭付近の基準点の位置のうち、少なくとも一方の情報を求めることを特徴とする請求項 1 記載の管理情報抽出装置。

【請求項 3】 前記計算手段は、前記表領域の外郭部分の情報として、該表領域の外郭付近の 2 つ以上の基準点の位置を求め、該 2 つ以上の基準点に対する相対的な位置情報に基づいて、前記管理情報の位置を算出することを特徴とする請求項 1 記載の管理情報抽出装置。

【請求項 4】 前記計算手段は、2 つの罫線の交点の位置情報、2 つの罫線の交差形状、入力画像に含まれる交点の数、および罫線により囲まれた特定の形状の矩形セルの頻度のうち、少なくとも 1 つ以上の情報を罫線構造の特徴として用いて、前記管理情報の位置を算出することを特徴とする請求項 1 記載の管理情報抽出装置。

【請求項 5】 前記計算手段は、罫線の種類が実線の場合と破線の場合とを区別して、前記罫線構造の特徴を求めることを特徴とする請求項 4 記載の管理情報抽出装置。

【請求項 6】 前記計算手段は、前記罫線の抽出の信頼度を罫線構造の特徴として用いて、前記管理情報の位置を算出することを特徴とする請求項 1 記載の管理情報抽出装置。

【請求項 7】 前記計算手段は、前記罫線上に並んでいる複数の交点間の 2 つ以上の間隔の比率を、罫線構造の特徴として用いて、前記管理情報の位置を算出することを特徴とする請求項 1 記載の管理情報抽出装置。

【請求項 8】 前記計算手段は、複数の罫線上で、前記複数の交点の並びを前記表領域の外郭付近から抽出し、前記間隔の比率を各罫線に対応する要素とする特徴ベクトルを求め、該特徴ベクトルを用いて前記表領域の外郭形状の特徴を表現することを特徴とする請求項 7 記載の管理情報抽出装置。

【請求項 9】 前記計算手段は、前記入力画像の外側から、左右上下の 4 方向のうち少なくとも 1 つ以上の方向で前記表領域の外郭形状の特徴を求め、該外郭形状の特徴を用いて前記管理情報の位置を算出することを特徴とする請求項 1 記載の管理情報抽出装置。

【請求項 1 0】 1 つ以上の表のフォームの罫線構造の特徴と、各表のフォーム内の管理情報の位置情報とを格納する辞書手段と、前記入力画像の罫線構造の特徴を前

記辞書手段に格納された罫線構造の特徴と照合する照合手段とをさらに備え、前記計算手段は、前記照合手段による照合結果に基づいて、前記辞書手段に格納された管理情報の位置情報を参照し、前記入力画像の管理情報の位置を算出することを特徴とする請求項 1 記載の管理情報抽出装置。

【請求項 1 1】 前記照合手段は、大分類用の罫線構造の特徴を用いて照合の候補となる表のフォームを絞り込み、詳細識別用の罫線構造の特徴を用いて照合を行い、前記入力画像に対応する表のフォームを決定することを特徴とする請求項 1 0 記載の管理情報抽出装置。

【請求項 1 2】 前記照合手段は、ダイナミック・プログラミング・マッチングにより、前記入力画像に対応する表のフォームを決定することを特徴とする請求項 1 1 記載の管理情報抽出装置。

【請求項 1 3】 前記辞書手段は、各表のフォーム内の前記管理情報の位置情報として、該管理情報を囲む矩形セルの位置情報を格納することを特徴とする請求項 1 0 記載の管理情報抽出装置。

【請求項 1 4】 前記辞書手段は、前記矩形セルの位置情報として、該矩形セルの 1 つ以上の頂点と、該矩形セルを含む表の 1 つ以上の頂点との間の差分ベクトルを格納することを特徴とする請求項 1 3 記載の管理情報抽出装置。

【請求項 1 5】 前記計算手段は、前記照合結果を用いて前記入力画像の表領域の安定な頂点を求め、該安定な頂点から前記差分ベクトルを用いて、前記入力画像の管理情報の位置を算出することを特徴とする請求項 1 4 記載の管理情報抽出装置。

【請求項 1 6】 前記辞書手段は、前記矩形セルの大きさをさらに格納し、前記計算手段は、前記矩形セルの大きさに対応する大きさを持ち、前記差分ベクトルにより指定される位置の近くに存在する矩形セルから、前記入力画像の管理情報の位置を算出することを特徴とする請求項 1 5 記載の管理情報抽出装置。

【請求項 1 7】 前記辞書手段は、各表の大きさをさらに格納し、前記計算手段は、前記入力画像の表領域の大きさと前記辞書手段内の対応する表の大きさから大きさの比率を計算し、該大きさの比率を用いて、前記入力画像の管理情報の位置を算出することを特徴とする請求項 1 3 記載の管理情報抽出装置。

【請求項 1 8】 前記照合手段は、前記入力画像から抽出された罫線と前記辞書手段の情報に含まれる罫線の間の対応可能な複数の組み合わせを求め、該複数の組み合わせの中から互いに両立する 2 つ以上の組み合わせのグループを抽出し、該グループに含まれる組み合わせの情報に基づいて、前記入力画像と表のフォームを照合することを特徴とする請求項 1 0 記載の管理情報抽出装置。

【請求項 1 9】 1 つ以上の表のフォームの罫線構造の特徴と、各表のフォーム内の管理情報の位置情報とを格

納する辞書手段と、

入力画像の罫線構造の特徴を前記辞書手段に格納された罫線構造の特徴と照合する照合手段と、

前記照合手段による照合結果に基づいて、前記辞書手段に格納された管理情報の位置情報を参照し、前記入力画像の管理情報を抽出する抽出手段と、

ユーザにより指定された管理情報の位置を前記辞書手段に登録するユーザ登録手段とを備えることを特徴とする管理情報抽出装置。

【請求項 2 0】 前記抽出手段が自動的に抽出した管理情報の位置を前記辞書手段に登録する自動登録手段をさらに備えることを特徴とする請求項 1 9 記載の管理情報抽出装置。

【請求項 2 1】 前記照合手段による照合の結果、前記入力画像が前記 1 つ以上の表のフォームのいずれにも対応しないとき、前記抽出手段は、前記入力画像から管理情報らしい部分を自動的に抽出することを特徴とする請求項 1 9 記載の管理情報抽出装置。

【請求項 2 2】 蓄積される画像の管理情報として、画像情報を格納する格納手段と、前記画像情報を検索する検索手段とを備えることを特徴とする画像蓄積装置。

【請求項 2 3】 任意の入力画像に含まれる表領域の外郭部分に対する罫線の相対的な位置情報に基づいて、該入力画像に含まれる管理情報の位置を算出する計算手段と、前記計算手段が算出した位置に基づいて、前記入力画像から前記画像情報を抽出して前記格納手段に登録する登録手段とをさらに備えることを特徴とする請求項 2 2 記載の画像蓄積装置。

【請求項 2 4】 前記画像情報と文字コードのうち少なくとも一方を選択する選択手段をさらに備え、前記格納手段は、選択された情報を前記管理情報として格納することを特徴とする請求項 2 2 記載の画像蓄積装置。

【請求項 2 5】 前記選択手段が前記管理情報に対応する文字コードを選択したとき、前記格納手段は、該管理情報に対する文字認識の信頼度に応じて、該管理情報の画像を前記画像情報として保存することを特徴とする請求項 2 4 記載の画像蓄積装置。

【請求項 2 6】 表のフォームの罫線情報を格納する格納手段と、

入力画像から抽出された罫線と前記格納手段の罫線情報に含まれる罫線の間の対応可能な複数の組み合わせを求め、該複数の組み合わせの中から互いに両立する 2 つ以上の組み合わせを含むグループを、他のグループの組み合わせが含まれないように抽出するグループ生成手段と、

抽出された 1 つ以上のグループに含まれる組み合わせの情報に基づいて、前記入力画像と表のフォームを照合する照合手段とを備えることを特徴とするフォーム識別装置。

【請求項 2 7】 前記 1 つ以上のグループの間で、罫線の対応関係が互いに矛盾しないグループの集合を抽出し、該グループの集合の中で、含まれる罫線の組み合わせの数が最も多い最適集合を求める集合決定手段をさらに備え、前記照合手段は、前記最適集合の各グループに含まれる組み合わせの情報に基づいて、前記入力画像と表のフォームを照合することを特徴とする請求項 2 6 記載のフォーム識別装置。

【請求項 2 8】 前記グループ生成手段は、前記入力画像の外郭部分の特徴量と各罫線の特徴量の相対値を、前記表のフォームの対応する特徴量の相対値と比較して、前記入力画像の罫線と前記表のフォームの罫線の間の対応可能性を判断し、対応の可能性がある罫線同士の組み合わせを生成することを特徴とする請求項 2 6 記載のフォーム識別装置。

【請求項 2 9】 前記グループ生成手段は、前記入力画像に含まれる罫線間の相対的關係を、前記表のフォームの罫線間の相対的關係と比較して、前記 2 つ以上の組み合わせが両立するかどうかを判定することを特徴とする請求項 2 6 記載のフォーム識別装置。

【請求項 3 0】 前記グループ生成手段は、前記入力画像の罫線を第 1 の方向に並べ、前記表のフォームの罫線を第 2 の方向に並べ、該入力画像の i 番目の罫線と該表のフォームの j 番目の罫線の組み合わせを、要素 (i, j) の位置のノードとするマッチングテーブルを作成するテーブル作成手段と、該マッチングテーブル上で、互いに両立する組み合わせに対応する 2 つのノードをパスで結ぶパス生成手段とを含むことを特徴とする請求項 2 6 記載のフォーム識別装置。

【請求項 3 1】 前記パス生成手段は、前記要素 (i, j) の位置のノードと両立する次のノードを、 $x > i$ かつ $y > j$ であるような要素 (x, y) の範囲で探索し、該次のノードが見つければ、該次のノードを基準として前記範囲と同様の探索範囲を設定することを特徴とする請求項 3 0 記載のフォーム識別装置。

【請求項 3 2】 表のフォームの罫線情報を格納する格納手段と、

入力画像の外郭部分に対する罫線の長さおよび位置の相対的な特徴を、前記表のフォームの罫線情報に含まれる罫線の対応する特徴と比較して、前記入力画像の罫線と前記表のフォームの罫線の間の対応可能性を判断し、対応の可能性がある罫線同士の組み合わせを生成する手段と、

前記組み合わせの情報に基づいて、前記入力画像と表のフォームを照合する照合手段とを備えることを特徴とするフォーム識別装置。

【請求項 3 3】 コンピュータのためのプログラムを記録した記録媒体であって、

任意の入力画像に含まれる表領域の外郭部分に対する罫線の相対的な位置情報に基づいて、該入力画像に含まれ

る管理情報の位置を算出する機能と、
算出された位置に基づいて、前記入力画像から前記管理
情報を抽出する機能とを前記コンピュータに実現させる
ためのプログラムを記録したコンピュータ読み取り可能
な記録媒体。

【請求項 3 4】 コンピュータのためのプログラムを記
録した記録媒体であって、
ユーザにより指定された 1 つ以上の表のフォームの管理
情報の位置を、位置情報としてあらかじめ登録する機能
と、
入力画像の罫線構造の特徴を、あらかじめ格納された前
記 1 つ以上の表のフォームの罫線構造の特徴と照合する
機能と、
照合結果に基づいて前記位置情報を参照し、前記入力画
像の管理情報を抽出する機能とを前記コンピュータに実
現させるためのプログラムを記録したコンピュータ読み
取り可能な記録媒体。

【請求項 3 5】 コンピュータのためのプログラムを記
録した記録媒体であって、
蓄積される画像の管理情報として、画像情報を格納する
機能と、
前記画像情報を検索する機能とを前記コンピュータに実
現させるためのプログラムを記録したコンピュータ読み
取り可能な記録媒体。

【請求項 3 6】 コンピュータのためのプログラムを記
録した記録媒体であって、
入力画像から抽出された罫線とあらかじめ格納された表
のフォームの罫線の間の対応可能な複数の組み合わせを
求める機能と、
前記複数の組み合わせの中から、互いに両立する 2 つ以
上の組み合わせを含むグループを、他のグループの組み
合わせが含まれないように抽出する機能と、抽出された
1 つ以上のグループに含まれる組み合わせの情報に基づ
いて、前記入力画像と表のフォームを照合する機能とを
前記コンピュータに実現させるためのプログラムを記録
したコンピュータ読み取り可能な記録媒体。

【請求項 3 7】 コンピュータのためのプログラムを記
録した記録媒体であって、
入力画像の外郭部分に対する罫線の長さおよび位置の相
対的な特徴を、前記表のフォームの罫線情報に含まれる
罫線の対応する特徴と比較して、前記入力画像の罫線と
前記表のフォームの罫線の間の対応可能性を判断し、対
応の可能性のある罫線同士の組み合わせを生成する機能
と、
前記組み合わせの情報に基づいて、前記入力画像と表の
フォームを照合する機能とを前記コンピュータに実現さ
せるためのプログラムを記録したコンピュータ読み取り
可能な記録媒体。

【請求項 3 8】 任意の入力画像に含まれる表領域の外
郭部分に対する罫線の相対的な位置情報に基づいて、該

入力画像に含まれる管理情報の位置を算出し、
算出された位置に基づいて、前記入力画像から前記管理
情報を抽出することを特徴とする管理情報抽出方法。

【請求項 3 9】 ユーザにより指定された 1 つ以上の表
のフォームの管理情報の位置を、位置情報としてあらか
じめ登録し、

入力画像の罫線構造の特徴を、あらかじめ格納された前
記 1 つ以上の表のフォームの罫線構造の特徴と照合し、
照合結果に基づいて前記位置情報を参照し、前記入力画
像の管理情報を抽出することを特徴とする管理情報抽出
方法。

【請求項 4 0】 蓄積される画像の管理情報として、画
像情報を登録し、
前記画像情報を検索することを特徴とする画像蓄積方
法。

【請求項 4 1】 入力画像から抽出された罫線とあらか
じめ格納された表のフォームの罫線の間の対応可能な複
数の組み合わせを求め、
前記複数の組み合わせの中から、互いに両立する 2 つ以
上の組み合わせを含むグループを、他のグループの組み
合わせが含まれないように抽出し、
抽出された 1 つ以上のグループに含まれる組み合わせの
情報に基づいて、前記入力画像と表のフォームを照合す
ることを特徴とするフォーム識別方法。

【請求項 4 2】 入力画像の外郭部分に対する罫線の長
さおよび位置の相対的な特徴を、前記表のフォームの罫
線情報に含まれる罫線の対応する特徴と比較して、前記
入力画像の罫線と前記表のフォームの罫線の間の対応可
能性を判断し、

対応の可能性のある罫線同士の組み合わせを生成し、
前記組み合わせの情報に基づいて、前記入力画像と表の
フォームを照合することを特徴とするフォーム識別方
法。

【発明の詳細な説明】

【0 0 0 1】

【発明の属する技術分野】本発明は、文書や図面をスキ
ャナ等の入力装置で画像データに変換し、その画像デー
タに管理情報を付加して蓄積するシステム、画像認識の
ために画像中の罫線構造を識別する装置、およびそれら
の処理を行う方法に関する。

【0 0 0 2】

【従来の技術】最近、従来より紙媒体で保存してきた情
報を電子化する動きが多く見られる。その中の 1 つに電
子ファイリングシステムがある。電子ファイリングシス
テムにおいては、紙文書がイメージスキャナ等の光電変
換装置で文書画像に変換され、それに検索のためのキー
ワード等の管理情報が付与されて、光ディスクやハード
ディスクに保存される。

【0 0 0 3】このような方法では、文書を画像データと
して保存するため、文書に記されているすべての文字を

文字認識技術でコード化してから保存する方法よりも、ディスク容量は多く必要となる。その反面、手軽で処理速度が速く、文字以外の絵や表などもそのまま保存できるメリットがある。しかし、保存された情報を検索するために、文書画像と共にキーワードや番号等の管理情報を付与しなければならない。従来のシステムはこのキーワード付けの手間がかかり、使いづらいシステムであった。

【0004】この使いづらさを解決するために、文書中にあるタイトル部分をキーワードとみなしてそれを自動的に抽出し、その部分を文字認識してコード化した結果を文書画像と共に保存する方法が考えられる。

【0005】現在の文字認識の処理速度は速くても数十文字/秒であり、A4の文書1枚を処理するには30秒から数分の処理時間がかかってしまう。したがって、タイトル抽出を高速化するためには、文書全体を文字認識するのではなく、まず画像上で必要なタイトル部分だけを抽出した後に、それを文字認識する方法が有力である。

【0006】一般文書を光電変換装置で読み込んで得られる文書画像から、文書のタイトル等の部分領域を抽出する先願の技術として、本出願人による「文書画像からのタイトル抽出装置および方法」(特願平7-341983)がある。このタイトル抽出装置の原理構成を図49に示す。

【0007】図49のタイトル抽出装置は、文字領域生成部1、文字列領域生成部2、およびタイトル抽出部3を備える。まず、文字領域生成部1は、スキャナ等で入力した文書画像から、画素の連結成分のラベル付け(ラベリング)により文字等の部分パターンを抽出する。そして、いくつかの部分パターンを統合することで、文字領域を抽出(生成)する。次に、文字列領域生成部2は、複数の文字領域を統合して、文字列領域を抽出(生成)し、タイトル抽出部3は、タイトルらしい文字列領域をタイトル領域として抽出する。

【0008】このとき、タイトル抽出部3は、タイトル領域である確からしさとして、文書の上部中央に位置している、本文と比べて文字サイズが大きい、下線が引かれている等の性質に着目する。そして、その確からしさを得点として各文字列領域に投票していき、最終的に得点の高い文字列領域の順番で、複数のタイトル領域の候補を算出する。表を含まない文書からは、以上の処理でタイトル領域を抽出できる。

【0009】これに対して、表を含む文書の場合、文字列領域生成部2が表中の文字列領域を抽出した後、タイトル抽出部3は、文字数の条件も考慮してタイトル領域を抽出する。例えば、「議題」、「名称」等、タイトルの存在を示唆する項目名の文字数は少なく、タイトル自身の文字列は「――について」というように文字数が多い可能性が高い。このことを利用して、隣り合う文字

列の文字数の関係から、タイトルらしい文字列を見つけることができる。

【0010】

【発明が解決しようとする課題】しかしながら、伝票等、罫線によってフォーマットが構造化された表形式文書は、現在大量に流通しており、一般には種々の表形式文書がある。このため、上述した先願の技術では、表内のタイトル抽出の精度が低いという問題がある。

【0011】例えば、タイトルが表内の中央部分や下端付近に存在する場合は、上部から優先的に文字列を抽出しただけでは、正しいタイトルが抽出結果に含まれない可能性がある。また、図50に示すように、承認欄11が表内の上部にあり、その中に「統括部長」、「部長」、「課長」、「担当者」等の余分な文字列が多く存在する場合も、これらの文字列が優先的に抽出され、正しいタイトルが抽出されにくい。

【0012】また、項目名12と、それと対になったタイトル13のように、タイトルが項目名の右ではなく下にある場合もある。この場合は、隣り合う文字列同士の文字数の情報だけでは、項目名とタイトルの位置関係を把握することができない。さらに、項目名が横書きだけでなく、縦書きの場合もあり、項目名の場所を正確に特定することは困難である。また、文書中に2つの表が含まれていて、タイトルが小さい方の表中のどこかに存在したりすることもある。

【0013】このように、表形式文書では種々のフォーマット(書式)があるため、タイトルとしての確からしさが文書によってばらつき、表内のタイトル抽出の精度が低くなってしまふ。入力された文書の画像の状態が悪ければ、抽出精度はさらに低くなる。

【0014】また、電子ファイリングシステムでは、抽出したタイトル領域を光学文字読み取り装置(optical character reader: OCR)で文字認識して文字コードを生成し、それを管理情報として画像に付加しておく。これにより、後に、データベース中の画像を文字コードで検索することができる。

【0015】この場合、タイトル領域の文字列がOCRで読み取りやすい文字であれば問題ないが、新聞、雑誌等によく見られるように、背景が模様(テクスチャ)になっていたり、文字がデザインフォントであったりする場合には、現状のOCRでは文字列を認識することができない。したがって、このような場合に、管理情報を画像に付加できないという問題がある。

【0016】本発明の課題は、種々の表形式文書を含む任意の画像から、その画像を管理するための適切な管理情報を抽出する管理情報抽出装置、その管理情報を用いて画像を蓄積する画像蓄積装置、およびそれらの方法を提供することである。

【0017】

【課題を解決するための手段】図1は、本発明の管理情

報抽出装置と画像蓄積装置を含む画像管理システムの原理図である。図1のシステムは、本発明の第1、第2、第3、および第4の原理を含み、ユーザ登録手段21、計算手段22、辞書手段23、照合手段24、抽出手段25、格納手段26、グループ生成手段27、および検索手段28を備える。

【0018】第1の原理において、計算手段22は、任意の入力画像に含まれる表領域の外郭部分に対する罫線の相対的な位置情報に基づいて、その入力画像に含まれる管理情報の位置を算出する。そして、抽出手段25は、計算手段22が算出した位置に基づいて、上記入力画像から上記管理情報を抽出する。

【0019】表領域の外郭部分の情報としては、例えば、表領域において縦横の基準となる大きさや、表領域の外郭付近に設けられた基準点の位置等が用いられる。計算手段22は、表領域内から抽出された各罫線の位置を上記基準点に対する相対的な位置情報として表し、管理情報を囲む罫線の位置情報から管理情報の位置を求める。抽出手段25は、その位置に対応する画像データを管理情報として抽出し、必要に応じて文字認識等を行

う。

【0020】表の外郭部分の複数の基準点あるいは複数の方向に対して、管理情報を囲む罫線の相対的な位置を求めることで、かすれやノイズ等のため入力画像の状態が悪い場合でも、精度良く管理情報を抽出できる。

【0021】また、第2の原理において、辞書手段23は、1つ以上の表のフォームの罫線構造の特徴と、各表のフォーム内の管理情報の位置情報とを格納する。照合手段24は、入力画像の罫線構造の特徴を辞書手段23に格納された罫線構造の特徴と照合する。抽出手段25は、照合手段24による照合結果に基づいて、辞書手段23に格納された管理情報の位置情報を参照し、上記入力画像の管理情報を抽出する。ユーザ登録手段21は、ユーザにより指定された管理情報の位置を辞書手段23に登録する。

【0022】表のフォームとは、表を構成する罫線のレイアウト構造を表し、辞書手段23は、その罫線構造の特徴とユーザ登録手段21により指定された管理情報の位置をあらかじめ格納しておく。そして、照合手段24は、入力画像の罫線構造の特徴と類似する特徴を持つ表のフォームを求め、抽出手段25は、その表のフォーム内で指定された位置から管理情報を抽出する。

【0023】このように、ユーザが希望する管理情報の位置をあらかじめ登録しておき、その位置に基づいて入力画像から管理情報を抽出することで、種々のフォームの画像が入力される場合でも、各画像から精度良く管理情報を抽出できる。

【0024】また、第3の原理において、格納手段26は、蓄積される画像の管理情報として、画像情報を格納する。そして、検索手段28は、上記画像情報を検索す

る。例えば、多数の画像を蓄積する電子ファイリング装置において、各画像から抽出された画像コードをその管理情報として格納手段26に格納しておく。そして、検索手段28は、例えば、テンプレートマッチングにより、与えられた画像コードを格納手段26内の画像コードと照合しながら、管理情報を検索する。

【0025】このように、管理情報の文字列を文字コードで格納/検索するだけでなく、画像そのもので格納/検索することもできるので、テキスト付き文字やデザインフォント、ロゴ等の正確に認識しにくい文字でも、管理情報として扱うことができる。

【0026】また、第4の原理において、格納手段26は、表のフォームの罫線情報を格納する。グループ生成手段27は、入力画像から抽出された罫線と格納手段26の罫線情報に含まれる罫線の間の対応可能な複数の組み合わせを求め、その複数の組み合わせの中から互いに両立する2つ以上の組み合わせを含むグループを、他のグループの組み合わせが含まれないように抽出する。照合手段24は、抽出された1つ以上のグループに含まれる組み合わせの情報に基づいて、上記入力画像と表のフォームを照合する。

【0027】グループ生成手段27は、格納手段26に格納された表のフォームを用いて入力画像のフォームを識別するために、入力画像の罫線と表のフォームの罫線の対応可能な組み合わせを求める。このとき、例えば、表全体に対する相対的な大きさや位置が類似している罫線同士が、対応可能な組み合わせとして取り出される。

【0028】そして、例えば、入力画像に含まれる罫線間の相対的關係を、表のフォームの罫線間の相対的關係と比較して、2つの組み合わせが両立するかどうかを判定し、互いに両立する組み合わせ同士のグループを生成する。このとき、すでに他のグループに属している組み合わせを含まないように、新しいグループを生成することで、両立性のチェック対象の数が削減され、処理が効率化される。

【0029】照合手段24は、例えば、こうして得られたグループの最適な集合に含まれる組み合わせの数が多く、入力画像と表のフォームの類似度が高いものとみなし、類似度の最も高い表のフォームを入力画像に対応するフォームと判断する。

【0030】これにより、入力画像のフォームを高速に識別することが可能になり、管理情報の抽出処理が効率化される。例えば、図1のユーザ登録手段21は、後述する図5における入力装置43に対応し、辞書手段23と格納手段26は外部記憶装置45に対応する。また、例えば、計算手段22、照合手段24、抽出手段25、グループ生成手段27、および検索手段28は、CPU41（中央処理装置）とメモリ42に対応する。

【0031】

【発明の実施の形態】以下、図面を参照しながら、本発

明の実施の形態を詳細に説明する。本発明においては、種々の表形式文書に対応できるように、既知の表形式の罫線のレイアウト構造を学習し、学習した情報を活用して未知の表形式文書からタイトル等を高精度に抽出する。このために、フォーム学習モードと運用モードを設ける。以下では、このレイアウト構造のことを、フォーマット構造またはフォームとも呼ぶことがある。

【0032】図2は、このような管理情報抽出処理の概要を示している。管理情報抽出装置は、まずフォーム学習時に、既知の表形式文書A、B等の罫線のレイアウト構造とユーザに指定させた正しいタイトル領域等の位置を学習する。そして、それらの情報を含むレイアウト辞書（フォーム辞書）31を作成しておく。

【0033】ユーザがタイトル位置を指定するモードとしては、文書A、B等のフォーム認識を伴わないユーザ登録モードと、フォーム認識を伴う自動学習モードとがある。これらの各モードにおける処理については、後述することにする。

【0034】次に、運用時には、管理情報抽出装置は、入力された未知の文書32から罫線のレイアウト構造を抽出し、レイアウト辞書31とのマッチングを行う。これにより、レイアウト辞書に記憶されているレイアウト構造と一致する表形式文書が識別される。ここでは、文書32のレイアウト構造は文書Aのそれと一致している。

【0035】そして、管理情報抽出装置は、対応する文書Aに指定されたタイトル位置の情報を参照して、入力文書32の文字列領域33からタイトルを高精度に抽出する。また、タイトルだけでなく、日付等の他のタグ領域を管理情報としてユーザに指定させることで、様々な文書の管理情報を高精度に抽出することができる。

【0036】運用時に、ユーザが文書をスキャナで入力する時に、管理情報が迅速に自動抽出されることが望まれるので、本発明では、ユーザのインタラクティブ性を重視した高速アルゴリズムを採用する。このアルゴリズムでは、対応するフォームを大分類、詳細分類の順に徐々に絞り込むことで、処理の高速化を図っており、フォーム学習時にも、これに対応する処理が行われる。

【0037】図3は、フォーム学習モードにおける処理の概要を示すフローチャートである。処理が開始されると、管理情報抽出装置は、まず、学習対象となる文書画像を入力し（ステップS1）、その罫線構造を抽出する（ステップS2）。そして、ユーザに管理情報の位置を問い合わせ、その位置を指定させる（ステップS3）。

【0038】次に、抽出された罫線構造から、実線と破線を区別しながら、大分類用の罫線の特徴を抽出し（ステップS4）、詳細識別用の外郭罫線の特徴を抽出する（ステップS5）。例えば、大分類用の特徴としては、データの変動に強い罫線構造の特徴を用い、詳細識別用の特徴としては、高速処理を考慮して、表の外郭形状に

関するCROSS RATIO（クロス比、複比）を用いる。

【0039】そして、抽出された罫線の特徴と指定された管理情報の位置を、レイアウト辞書31に格納して（ステップS6）、処理を終了する。格納された情報は運用モードにおいて参照され、未知文書から管理情報を抽出するために利用される。

【0040】図4は、運用モードにおける処理の概要を示すフローチャートである。処理が開始されると、管理情報抽出装置は、まず、処理対象となる文書画像を入力し（ステップS11）、その罫線構造を抽出する（ステップS12）。

【0041】次に、抽出された罫線構造から、大分類用の罫線の特徴を抽出し（ステップS13）、レイアウト辞書31の対応する情報と比較して、罫線構造の大分類を行う（ステップS14）。これにより、レイアウト辞書31の罫線構造のうち、入力文書のそれと一致する可能性のあるものが、候補として抽出される。

【0042】次に、罫線構造から詳細識別用の外郭罫線の特徴を抽出し（ステップS15）、大分類で抽出された候補の対応する情報と比較して、罫線構造の詳細識別を行う（ステップS16）。ここでは、例えば、CROSS RATIOの1次元マッチング処理を行って、入力文書に対応する候補を特定する。

【0043】そして、その候補のフォーム上で指定された管理情報の位置に基づいて、入力文書画像内の管理情報の位置を算出し（ステップS17）、処理を終了する。こうして、既知の文書中でユーザが指定した位置情報を用いて、入力文書画像から、管理情報を高精度に抽出することができる。また、運用時におけるフォームの照合処理を大分類と詳細識別の2段階で行うので、詳細識別の候補が絞り込まれ、抽出処理が高速化される。

【0044】本実施形態の管理情報抽出装置は、例えば図5に示すような情報処理装置（コンピュータ）により実現される。図5の情報処理装置は、CPU41、メモリ42、入力装置43、出力装置44、外部記憶装置45、媒体駆動装置46、ネットワーク接続装置47、および光電変換装置48を備え、それらの各装置はバス49により互いに結合されている。

【0045】CPU41は、メモリ42を利用しながらプログラムを実行して、図3、4に示した各処理を行う。メモリ42としては、例えばROM（read only memory）、RAM（random access memory）等が用いられる。ROMには、上述のプログラムが格納され、RAMには、レイアウト辞書31等の必要なデータが一時的に保持される。

【0046】入力装置43は、例えばキーボード、ポインティングデバイス等に相当し、ユーザからの要求や指示の入力に用いられる。また、出力装置44は、表示装置（ディスプレイ）やプリンタ等に相当し、ユーザへの問い合わせや処理結果等の出力に用いられる。

【0047】外部記憶装置45は、例えば、磁気ディスク装置、光ディスク装置、光磁気ディスク装置等である。この外部記憶装置45に、上述のプログラムとデータを保存しておき、必要に応じて、それらをメモリ42にロードして使用することができる。また、外部記憶装置45は、画像やレイアウト辞書31を保存するデータベースとしても使用される。

【0048】媒体駆動装置46は、可搬記録媒体50を駆動し、その記憶内容にアクセスすることができる。可搬記録媒体50としては、メモリカード、フロッピーディスク、CD-ROM (compact disk read only memory)、光ディスク、光磁気ディスク等、任意のコンピュータ読み取り可能な記録媒体を使用することができる。この可搬記録媒体50に、上述のプログラムとデータを格納しておき、必要に応じて、それらをメモリ42にロードして使用することができる。

【0049】ネットワーク接続装置47は、LAN (local area network) 等の任意の通信ネットワークに接続され、通信に伴うデータ変換等を行って、外部の情報提供者のデータベース50'等と通信する。これにより、管理情報抽出装置は、必要に応じて、上述のプログラムとデータをデータベース50'からネットワークを介して受け取り、それらをメモリ42にロードして使用することができる。

【0050】また、光電変換装置48は、例えばイメージスキャナであり、処理対象となる文書や図面等の画像を入力する。次に、図6から図16までを参照しながら、フォーム学習時に行われる各処理について説明する。

【0051】図6は、フォーム学習時の処理の詳細を示すフローチャートである。図6において、図3と対応する処理ステップには、図3と同じ符号が用いられている。まず、ステップS2の罫線構造抽出処理では、管理情報抽出装置は、図7に示すように、入力された文書画像から縦／横の破線を抽出し(ステップS2-1)、縦／横の実線を抽出し(ステップS2-2)、さらに、それらの縦横の罫線で囲まれた矩形セル(矩形領域)を抽出する(ステップS2-3)。

【0052】罫線や矩形セルの抽出には、本出願人による「画像抽出装置」(特開平7-28937)や「枠抽出装置及び矩形抽出装置」(特願平7-203259)等の先願の技術が用いられる。これらの技術によれば、帳票内の罫線枠の位置等の情報入力が必要とせず、その画像から枠の抽出や除去を行うことが可能である。以下に、罫線構造抽出処理の概要を示す。

【0053】(1) 細線化: マスク処理によって、縦横の線分を細線化し、文字と枠の太さの差をなくす。

(2) 線分抽出: 隣接投影法を用いて、比較的長い線分の抽出を行う。隣接投影とは、注目する行または列に含まれる画素の投影値に、その周囲の行または列の投影値

を足し合せた結果を、注目する行または列の最終的な投影値とする方法である。この投影法によれば、特定の行または列の周囲の画素分布を大局的に捉えることができる。

【0054】(3) 直線抽出: 抽出された線分を順に探索し、線分と線分の間在一定長さ以上の途切れがないかどうかを検査する。間にそのような途切れがない線分同士を順に統合していき、長い直線を抽出する。

【0055】(4) 直線統合: 抽出された直線同士を再度統合する。かすれにより2つ以上の部分に分離された直線を、1つの直線に統合する。

(5) 直線伸張: 規則的な帳票であると分かっているとに限り、かすれにより短くなった直線の伸張処理を行い、本来の長さに修復する。

【0056】(6) 横枠決定: 「枠抽出装置及び矩形抽出装置」(特願平7-203259)に示された規則に従って、表の上部から順に、2本1組で1行の記入欄を構成する横直線を、横枠線として2本ずつ抽出する。

【0057】(7) 縦枠決定: 縦枠線は、上述の記入欄1行毎に決定する。注目している行を構成する2本の横枠線に、上下端がともに達している縦直線を、その行における縦枠線とする。

【0058】(8) 矩形セル抽出: 2本の横枠線および2本の縦枠線で囲まれた矩形セルを、文字領域として抽出する。次に、ステップS3の管理情報位置指定処理では、管理情報抽出装置は、図8に示すように、入力された文書画像をディスプレイの画面に表示して、例えば、タイトルを表す文字列のどこかを、ユーザにマウスでポインティングしてもらう。そして、そのポインティング位置を含む矩形セル51の位置情報を記憶する。

【0059】矩形セル51の位置情報としては、表の周辺上の任意の交点(外郭点)を基準点とし、その点から矩形セル51の位置までのベクトルの情報が記憶される。例えば、左上頂点52、左下頂点53、右上頂点54、右下頂点55をベクトルの始点とした場合、それらの各頂点から、それぞれ、矩形セル51の左上頂点56、左下頂点57、右上頂点58、右下頂点59までの差分ベクトルA、B、C、Dのデータが記憶される。また、同時に、表の縦、横のサイズh0、w0、および、矩形セルの縦、横のサイズH1、W1も記憶しておく。

【0060】次に、ステップS4の大分類用罫線特徴抽出処理では、管理情報抽出装置は、まず、横罫線と縦罫線の交点の数を数える(ステップS4-1)。次に、各交点における交差形状を抽出し、その頻度分布を求める(ステップS4-2)。交差形状は、交点から上下左右の向きに罫線が存在するかどうかと、存在する罫線の種類を表すコード(K1、K2、K3、K4)で表現される。

【0061】ここで、要素K1は交点の上側の罫線に対応し、要素K2は交点の下側の罫線に対応し、要素K3

は交点の左側の野線に対応し、要素K 4は交点の右側の野線に対応する。各要素の値は、野線が存在しないとき0となり、実線が存在するとき1となり、破線が存在するとき2となる。

【0062】例えば、図9のような交点の交差形状は(1, 1, 1, 1)と表され、図10のような交点の交差形状は(1, 1, 1, 0)と表され、図11のような交点の交差形状は(0, 2, 2, 2)と表され、図12のような交点の交差形状は(1, 1, 2, 2)と表される。このように、(K 1, K 2, K 3, K 4)の各要素は3通りの値を取り得るため、可能なコードの総数は3⁴(=81)となる。ステップS 4-2では、81通りの各コード毎に出現回数(頻度)を求めて、それを記憶する。

【0063】次に、各矩形セルの縦横比を計算し、その頻度分布を矩形セルの頻度分布として算出する(ステップS 4-3)。矩形セルの縦の長さをH 1、横の長さをW 1とすると、縦横比はW 1/H 1のように求められる。縦横比の頻度分布は、例えば、W 1/H 1の値を0から0.5刻みで増やしていき、各値に対応する縦横比を持つ矩形セルの数をカウントすることで求められる。このとき、あるしきい値(例えば、10)を超える矩形セルについては、それらの数をまとめてカウントする。

【0064】次に、ステップS 5の詳細識別用外郭野線特徴抽出処理では、管理情報抽出装置は、まず、交点が並んでいる各行または列内で、左右または上下方向の外側(外郭)から、4つの交点から成る交点列を取り出す。

【0065】例えば、図13に示すような野線構造の場合は、2行目の左から順に4つの交点を取り出すと、交点61、62、63、64が取り出され、右から順に4つの交点を取り出すと、交点65、64、63、62が取り出される。また、3列目の上から順に4つの交点を取り出すと、交点66、63、67、68が取り出され、下から順に4つの交点を取り出すと、交点70、69、68、67が取り出される。

【0066】そして、取り出された交点列に関する1次元Projective Invariants(射影不変量)のCROSS RATIOを計算する。例えば、図14に示すように、4つの交点X 1、X 2、X 3、X 4から成る交点列が取り出されたとすると、そのCROSS RATIOは、次式で与えられる。

【0067】

【数1】

$$\text{CROSS RATIO} = \frac{|X1-X2| \cdot |X3-X4|}{|X1-X3| \cdot |X2-X4|} \quad (1)$$

【0068】ここで、|X i - X j|は、交点X iと交点X jの間隔の長さ(距離)を表す(i, j=1, 2, 3, 4)。(1)式のCROSS RATIOは、例えば、図15のようなフローチャートに従って計算される。CROSS RA

TIO 算出処理が開始されると、管理情報抽出装置は、まず、4つの交点X 1、X 2、X 3、X 4の座標データを入力する(ステップS 21)。

【0069】次に、交点X 1とX 2の距離を求めて変数aに入力し(ステップS 22)、交点X 3とX 4の距離を求めて変数bに入力し(ステップS 23)、交点X 1とX 3の距離を求めて変数cに入力し(ステップS 24)、交点X 2とX 4の距離を求めて変数dを入力する(ステップS 25)。そして、a b/c dを計算して、その値をCROSS RATIOとして記憶し(ステップS 26)、処理を終了する。

【0070】このようにして、すべての交点列のCROSS RATIOを算出することで、表の外郭付近に位置する交点の並びの特徴が定量化される。これにより、図16に示すように、表の外郭形状の2次元的な特徴が1次元の数値の並びで表現される。以下では、このCROSS RATIOの値の並びをCROSS RATIO列と呼ぶことにする。

【0071】図16において、右のCROSS RATIO列R[1], R[2], R[3], ..., R[n]は、各行の右端部分の特徴を表すCROSS RATIOに対応し、左のCROSS RATIO列L[1], L[2], L[3], ..., L[m]は、各行の左端部分の特徴を表すCROSS RATIOに対応する。また、上のCROSS RATIO列U[1], U[2], U[3], ..., U[w]は、各列の上端部分の特徴を表すCROSS RATIOに対応し、下のCROSS RATIO列D[1], D[2], D[3], ..., D[v]は、各行の下端部分の特徴を表すCROSS RATIOに対応する。

【0072】一般には、表の左右両端部分における野線構造は対称ではなく、また、画像の一部がかすれたり、潰れたりしている場合もあるので、nとmは必ずしも一致するとは限らない。同様の理由で、wとvは必ずしも一致するとは限らない。

【0073】これらの4つの方向のCROSS RATIO列を1つに繋げれば、各CROSS RATIOの値を要素とする特徴ベクトル(R[1], ..., R[n], L[1], ..., L[m], U[1], ..., U[w], D[1], ..., D[v])が生成される。

【0074】ところで、ここでは、詳細識別用の外郭野線の特徴として、4つの交点の間隔の比率を用いているが、その代わりに、2つ以上の任意の個数の交点の間隔の比率を用いてもよい。その場合も、その比率を1次元に並べることで、表の外郭形状の特徴が表現される。

【0075】そして、ステップS 6の処理では、管理情報抽出装置は、ステップS 3で指定された管理情報の位置とステップS 4、S 5で求めた野線の特徴とを、その表形式文書の識別情報(フォーム情報)として、レイアウト辞書31に格納する。

【0076】次に、図17から図22までを参照しながら、運用時に行われる各処理について説明する。図17

は、フォーム学習時の処理の詳細を示すフローチャートである。図 17 において、図 4 と対応する処理ステップには、図 4 と同じ符号が用いられている。まず、ステップ S 12 の罫線構造抽出処理では、管理情報抽出装置は、フォーム学習時のステップ S 2 の処理と同様にし、入力された文書画像から縦／横の破線を抽出し（ステップ S 12-1）、縦／横の実線を抽出し（ステップ S 12-2）、それらの縦横の罫線で囲まれた矩形セルを抽出する（ステップ S 12-3）。

【0077】また、ステップ S 13 の大分類用罫線特徴抽出処理では、管理情報抽出装置は、フォーム学習時のステップ S 4 の処理と同様にし、横罫線と縦罫線の交点の数を数え（ステップ S 13-1）、各交点における交差形状の頻度分布を求め（ステップ S 13-2）、各矩形セルの縦横比の頻度分布を算出する（ステップ S 13-3）。

【0078】次に、ステップ S 14 の大分類処理では、管理情報抽出装置は、交点数、交差形状の頻度分布、および矩形セルの縦横比の頻度分布を用いて、レイアウト辞書 31 内の多数の表のフォーム情報との照合を行い、該当する表の候補の数を絞り込む。ここでは、画像のかすれ、潰れ等による罫線構造の変動も考慮して、交点数、交差形状の頻度、および矩形セルの縦横比の頻度の各特徴に対して適当なしきい値を設定し、辞書 31 のフォーム情報が一定の許容範囲内で入力画像の情報と一致すれば、それを該当する表の候補とする。

【0079】例えば、入力された文書画像の交点数を K_i とし、辞書 31 に格納されたフォーム t の交点数を K_t とすると、それらの差の絶対値 $|K_i - K_t|$ がしきい値 TH_k 以内であれば、フォーム t を候補とする。このように、入力画像と辞書 31 内のフォーム情報の各要素の差が、すべてそれぞれのしきい値以内に収まれば、そのフォームを入力文書に該当するフォームの候補とする。

【0080】交点数、交点形状、矩形セルのサイズの頻度分布等の特徴は、一般に、画像の変動に強く、これらを用いることで、かすれや潰れのある文書画像でも高精度な照合を行うことができる。

【0081】次に、ステップ S 15 の詳細識別用外郭罫線特徴抽出処理では、管理情報抽出装置は、フォーム学習時のステップ S 5 の処理と同様にし、4 方向からの 1 次元 Projective Invariants の CROSS RATIO を計算する。

$$E[i, k] = \min \{ E[i-1, k] + d[i, k], \\ E[i-1, k-1] + \lambda * d[i, k], \\ E[i, k-1] + d[i, k] \} \quad (2)$$

ここで、 $E[i, k]$ は、CROSS RATIO 列の一部分である $(R[1], \dots, R[i])$ を $(R'[1], \dots, R'[k])$ に対応付けたときの誤差の累積の最小値を表す。したがって、計算途中における

【0082】次に、ステップ S 16 の詳細識別処理では、管理情報抽出装置は、大分類で絞り込まれた表候補だけに限って、CROSS RATIO 列の照合を行う。この処理では、入力フォームと学習済みフォームの間で、4 方向それぞれ別々に CROSS RATIO 列の対応付けを行う。ここでは、対象とするフォームが表構造なので、各列と各行で罫線の順序関係が逆転することではなく、かすれ等により罫線が部分的に消失することだけを考慮し、DP (Dynamic Programming) マッチングを採用することにする。

【0083】DP マッチングは、音声等の時系列データのマッチング方法としてよく知られており、その詳細は、例えば、舟久保登著「パターン認識」（共立出版）p. 62-p. 67 に記述されている。この方法では、2 つのデータの集合を照合する際に、データの局所的な特徴に関して類似度を与え、類似度を用いて全体的な対応付けの良否を表す評価関数を定義する。そして、最も良い評価関数の値を与えるデータの対応関係を求める。

【0084】図 18 は、このような DP マッチングを用いた右の CROSS RATIO 列の照合処理を示している。図 18 においては、入力フォームの右の CROSS RATIO 列 $R[1], R[2], R[3], \dots, R[n]$ が、辞書 31 内の学習済みフォームの右の CROSS RATIO 列 $R'[1], R'[2], R'[3], \dots, R'[n']$ と対応付けられている。

【0085】また、この照合処理においては、罫線抽出の信頼度を考慮して、信頼度の高い罫線から求めた交点列の CROSS RATIO と、そうでない CROSS RATIO とで、評価関数における対応付けの重みを変えておく。例えば、信頼度の高い罫線から求めた CROSS RATIO の類似度には、より大きな重みが付加される。

【0086】図 19 は、DP マッチングによる右の CROSS RATIO 列の照合処理の例を示すフローチャートである。処理が開始されると、管理情報抽出装置は、まず、入力フォームの右の CROSS RATIO 列を配列 $R[i]$ ($i=1, \dots, n$) に格納し、学習済みフォームの右の CROSS RATIO 列を配列 $R'[k]$ ($k=1, \dots, n'$) に格納する（ステップ S 31）。

【0087】次に、誤差配列 $E[i, k]$ を初期化し（ステップ S 32）、 $i=1, \dots, n, k=1, \dots, n'$ について、次のような漸化式の計算を行う（ステップ S 33）。

【0088】

累積誤差を評価関数として用いた場合、 $E[i, k]$ はその最小値を与えていることになる。また、 $d[i, k]$ は、 $R[i]$ と $R'[k]$ を対応付けたときの誤差を表し、例えば、次式により与えられる。

【0089】

$$d[i, k] = |R[i] - R'[k]| \quad (3)$$

また、 λ は $d[i, k]$ に対する重みを表し、 $min\{\}$ は $\{\}$ 内の要素のうちの最小値を表す。

【0090】次に、 $E[n, n']$ の経路を算出し（ステップS34）、それをCROSS RATIO 列（ $R[1], \dots, R[n]$ ）と（ $R'[1], \dots, R'[n']$ ）の対応付けの結果として格納して（ステップS35）、処理を終了する。こうして、評価関数が最小となるような各CROSS RATIO 同士の対応関係が求められる。左、上、下のCROSS RATIO 列の照合処理についても同様である。

【0091】ステップS16では、このような1次元のDPマッチングを、大分類で得られたすべての学習済みフォームについて行い、それらのうちで評価関数が最小（最良）となるものを、入力フォームに該当するフォームと判断する。このように、詳細識別において、表構造の外側の形状（外郭形状）の特徴を用い、しかも1次元マッチングによる識別を行うので、高速処理が可能となる。

【0092】そして、ステップS17の管理情報位置算出処理では、管理情報抽出装置は、レイアウト辞書31を参照して、詳細識別で特定された学習済みフォームの位置情報を取り出し、それを用いて入力画像から管理情報を抽出する。

【0093】この処理では、まず、上述のDPマッチングにおけるCROSS RATIO 列の対応付け結果を用いて、表の各行各列の両端に位置する交点（端点）のマッチング度を調べ、端点が安定かどうかを判定する。ここで、端点のマッチング度とは、入力フォームのあるCROSS RATIO 0 と学習済みフォームのあるCROSS RATIO の対応付けの確からしさを意味する。

【0094】例えば、図18においては、 $R[1]$ と $R'[1]$ は一意的（1対1）に対応しているため、1行目の右端の端点は安定であると判定される。また、 $R[3]$ と $R'[4]$ も1対1に対応しているため、該当する行の右端の端点は安定である。しかし、 $R[2]$ は $R'[2]$ と $R'[3]$ の両方に対応付けられており、対応が一意的ではないため、該当する行の右端の端点は安定ではないと判定される。このようにして、表の左

上、左下、右上、右下の各頂点に最も近い安定した端点をそれぞれ求め、これらを安定な外郭点とする。

【0095】次に、安定な外郭点を基準にして、入力フォームの表と学習済みフォームの表の縦、横のサイズ h_0, w_0 を求め、それらを比較して、学習済みフォームの表に対する入力フォームの表の縦横の相対的な比率を求める。そして、図8に示した差分ベクトルA、B、C、D、および矩形セルの縦、横のサイズ $H1, W1$ を用いて、管理情報の位置を算出する。

【0096】上述の比率は、入力フォームの表の学習済

みフォームの表に対する拡大率または縮小率を表し、それらの表の間で、サイズの変動を正規化するために用いられる。

【0097】例えば、図8に示した表に対する入力フォームの表の縦横の比率がともに α である場合、差分ベクトルA、B、C、Dをそれぞれ α 倍する。そして、入力フォームの表内で、左上の安定な外郭点を始点として、差分ベクトルAの α 倍のベクトルを用いて、管理情報を含む矩形セルの左上頂点のおよその位置を求める。同様に、右上、左下、右下の安定な外郭点を始点として、差分ベクトルB、C、Dの α 倍のベクトルを用いれば、矩形セルの右上頂点、左下頂点、右下頂点のおよその位置が求められる。

【0098】次に、こうして求めた位置の近くに存在する矩形セルであって、縦、横のサイズがそれぞれ $H1 * \alpha, W1 * \alpha$ に類似している矩形セルを見つける。そして、その矩形セル内部にある文字列等のデータを、所望の管理情報として取り出す。

【0099】図20、21、22は、管理情報位置算出処理の例を示すフローチャートである。処理が開始されると、管理情報抽出装置は、まず、DPマッチングにおける4方向のCROSS RATIO 列の対応付け結果を入力する（ステップS41）。

【0100】ここでは、右のCROSS RATIO 列（ $R[1], \dots, R[n]$ ）と（ $R'[1], \dots, R'[n']$ ）の対応付けの結果と、左のCROSS RATIO 列（ $L[1], \dots, L[m]$ ）と（ $L'[1], \dots, L'[m']$ ）の対応付けの結果と、上のCROSS RATIO 列（ $U[1], \dots, U[w]$ ）と（ $U'[1], \dots, U'[w']$ ）の対応付けの結果と、下のCROSS RATIO 列（ $D[1], \dots, D[v]$ ）と（ $D'[1], \dots, D'[v']$ ）の対応付けの結果とが入力される。

【0101】次に、これらのデータから入力フォームの安定な端点を算出し、これらを安定な外郭点の候補とする（ステップS42）。これらの候補に対応するCROSS RATIO を、それぞれ、 $R[nmin], R[nmax], L[mmmin], L[mmmax], U[wmin], U[wmax], D[vmin], D[vmax]$ と書くことにする。

【0102】ここで、 $nmin$ は、表の右端の安定な端点のうち最も上（y座標が最小）にある点の行番号を表し、 $nmax$ は、表の右端の安定な端点のうち最も下（y座標が最大）にある点の行番号を表す。また、 $mmmin$ は、表の左端の安定な端点のうち最も上にある点の行番号を表し、 $mmmax$ は、表の左端の安定な端点のうち最も下にある点の行番号を表す。

【0103】また、 $wmin$ は、表の上端の安定な端点

のうち最も左 (x座標が最小) にある点の列番号を表し、wmaxは、表の上端の安定な端点のうち最も右 (x座標が最大) にある点の列番号を表す。また、vminは、表の下端の安定な端点のうち最も左にある点の列番号を表し、vmaxは、表の下端の安定な端点のうち最も右にある点の列番号を表す。

【0104】次に、得られた候補のデータを用いて、安定な外郭点の位置を算出する (ステップS43)。ここでは、各候補のx座標、y座標の最大値および最小値を求めて、それらの値を安定な外郭点の座標成分とする。

【0105】図20において、例えば、XMIN {R [nmin], R [nmax], L [mmmin], L [mmaax], U [wmin], U [wmax], D [vmin], D [vmax]} は、{} 内の各CROSS RATIO の値に対応する端点のx座標のうちで最小のものを表す。同様に、XMAX {} は各端点のx座標の最大値を表し、YMIN {} は各端点のy座標の最小値を表し、YMAX {} は各端点のy座標の最大値を表す。

【0106】以下では、これらの値XMIN {}, XMAX {}, YMIN {}, YMAX {} を、簡単のため、それぞれ、XMIN、XMAX、YMIN、YMAXと書くことにする。このとき、左上、右上、左下、右下の安定な外郭点の座標は、それぞれ、(XMIN, YMIN)、(XMAX, YMIN)、(XMIN, YMAX)、(XMAX, YMAX) で与えられる。

【0107】次に、学習済みフォームである辞書フォームの安定な端点を算出し、これらを安定な外郭点の候補とする (図21、ステップS44)。これらの候補に対応するCROSS RATIO を、それぞれ、R' [nmi

$$w0 = XMAX - XMIN$$

$$h0 = YMAX - YMIN$$

また、ステップS45で得られた安定な外郭点の座標情報を用いて、次式により辞書フォームの縦、横のサイズ

$$w0' = XMAX' - XMIN'$$

$$h0' = YMAX' - YMIN'$$

そして、縦、横のサイズh0、w0、h0'、w0' から、次式により入力フォームと辞書フォームの大きさの比率 (拡大率または縮小率) Sw、Shを算出する (ス

$$Sw = w0 / w0'$$

$$Sh = h0 / h0'$$

次に、辞書31から、辞書フォームの表の安定な外郭点を始点とする差分ベクトルの成分の大きさを、管理情報の位置を表す相対座標値として取り出す (ステップS49)。この場合、辞書フォームのCROSS RATIO R'

[1], ..., R' [n'], L' [1], ..., L' [m'], U' [1], ..., U' [w'], D' [1], ..., D' [v'] のそれぞれに対応する外郭点のうち、各頂点に近い複数の外郭点からの差分ベクトルが、あらかじめ辞書31に位置情報として格納されているものとする。

n'], R' [nmax'], L' [mmmin'], L' [mmaax'], U' [wmin'], U' [wmax'], D' [vmin'], D' [vmax'] と書くことにする。

【0108】ここで、nmin', nmax', mmmin', mmaax', wmin', wmax', vmin', vmax' の意味については、上述のnmin、nmax、mmmin、mmaax、wmin、wmax、vmin、vmaxと同様である。

【0109】次に、得られた候補のデータを用いて、ステップS43と同様に、辞書フォームの安定な外郭点の位置を算出する (ステップS45)。図21において、XMIN' {}, XMAX' {}, YMIN' {}, YMAX' {} の意味については、上述のXMIN {}, XMAX {}, YMIN {}, YMAX {} と同様である。

【0110】以下では、これらの値XMIN' {}, XMAX' {}, YMIN' {}, YMAX' {} を、簡単のため、それぞれ、XMIN'、XMAX'、YMIN'、YMAX' と書くことにする。このとき、左上、右上、左下、右下の安定な外郭点の座標は、それぞれ、(XMIN', YMIN')、(XMAX', YMIN')、(XMIN', YMAX')、(XMAX', YMAX') で与えられる。

【0111】次に、ステップS43で得られた安定な外郭点の座標情報を用いて、次式により入力フォームの縦、横のサイズh0、w0を算出する (図22、ステップS46)。

【0112】

$$(4)$$

$$(5)$$

h0'、w0' を算出する (ステップS47)。

【0113】

$$(6)$$

$$(7)$$

ステップS48)。

【0114】

$$(8)$$

$$(9)$$

【0115】そして、左上、右上、左下、右下の安定な外郭点から相対座標値を、それぞれ、(fxmin1, fymin1)、(fxmax1, fymin2)、(fxmin2, fymax1)、(fxmax2, fymax2) とする。

【0116】次に、これらの相対座標値と、入力フォームと辞書フォームの大きさの比率Sw、Shを用いて、入力フォーム内における管理情報の位置の概算を行う (ステップS50)。ここでは、管理情報の位置の候補として、次のような座標値を持つ4つの点が求められ

る。

【0117】 $(XMIN + Sw * fxmin1, YMIN + Sh * fymin1)$

$(XMAX - Sw * fxmax1, YMIN + Sh * fymin2)$

$(XMIN + Sw * fxmin2, YMAX - Sh * fymax1)$

$(XMAX - Sw * fxmax2, YMAX - Sh * fymax2)$

次に、これらの候補位置を含む入力フォームの矩形セルを抽出する(ステップS51)。そして、その縦のサイズが、辞書フォームで指定されている矩形セルの縦のサイズH1のSh倍に近く、横のサイズが、指定されている矩形セルの横のサイズW1のSw倍に近い場合に、その矩形セルを管理情報を含むセルと判定する。

【0118】そして、その矩形セル内の文字列等の画像データを管理情報として出力し(ステップS52)、処理を終了する。こうして、詳細識別の結果を用いて、入力画像から管理情報が抽出される。

【0119】ここでは、辞書31は、辞書フォームのCR OSS RATIO に対応する複数の外郭点の一部を始点とする差分ベクトルを格納するものとしたが、すべての外郭点からの差分ベクトルをあらかじめ格納しておく構成にしてもよい。こうしておくことで、表の頂点の近くの外郭点だけでなく、周辺上の任意の外郭点を安定な外郭点として選ぶことが可能になる。

【0120】また、安定な外郭点を4つ抽出する必要は必ずしもなく、処理を高速化するために、いずれか1つの安定な外郭点を基準点として、その位置からの相対座標値で管理情報の位置を求めてもよい。一般に、処理に用いる安定な外郭点の数は任意である。

【0121】さらに、ステップS51では、4つの候補位置を含む矩形セルを抽出するものとしたが、代わりに、1つ以上の候補位置を含む矩形セルを抽出してもよく、また、1つ以上の候補位置との距離が一定値以内であるような矩形セルを抽出してもよい。

【0122】上述した管理情報抽出処理によれば、入力文書のフォームと管理情報の位置が自動的に学習され、レイアウト辞書31に格納される。また、これらの情報を利用することで、種々の表形式文書に対応することができ、管理情報の位置が高精度に算出される。

【0123】次に、図6のステップS3における管理情報位置の指定方法について、さらに詳しく説明する。本実施形態では、ユーザによる管理情報位置の指定方法として、ユーザにその位置を明示的に指定させるユーザ登録モードと、管理情報の候補を自動的に抽出する自動学習モードの2通りを用意する。

【0124】ユーザ登録モードでは、管理情報抽出装置は、図8に示したように、表を構成する多数の矩形セルの中から管理情報の位置を、直接ユーザに指定させる。

例えば、設計図面等の同じフォームの文書が大量にある場合に、最初の1枚目で管理情報の位置を指定しておけば、2枚目からはその位置の情報のみを読み取ればよく、オートドキュメントフィードを使ったバッチ入力が可能になる。

【0125】また、自動学習モードでは、先願の「文書画像からのタイトル抽出装置および方法」(特願平7-341983)に記載されたタイトル抽出技術を用いて、管理情報の候補となる複数の領域を抽出し、それらの中からユーザが選択した領域の位置を自動学習して、次回からそれを1位の候補とする。もし、ユーザがそれらの候補のいずれも選択せずに、新たに任意の位置を指定した場合には、ユーザのインタラクティブな操作からその位置の情報を自動的に取得する。

【0126】あるいはまた、先願のタイトル抽出技術をユーザ登録モードにも適用して、複数の候補の中から管理情報を指定させることも可能である。この場合、自動学習モードでは、まず、図4の処理によりフォームの認識/識別を行い、入力画像が辞書31のフォームと一致するかどうかを調べる。そして、いずれかの辞書フォームと一致すれば、その位置情報を取り出してユーザに提示し、いずれの辞書フォームとも一致しなければ、先願のタイトル抽出技術により管理情報の候補を抽出する。

【0127】図23は、このような2つのモードを備えた管理情報抽出処理を示している。図23のユーザ登録モードにおいては、管理情報抽出装置は、まず、表形式文書の入力画像71から、先願の表内タイトル抽出処理を用いて、管理情報の複数の候補を抽出する。

【0128】図24は、この表内管理情報抽出処理のフローチャートである。処理が開始されると、管理情報抽出装置は、文書71を読み込み、文書画像としてメモリに格納する(ステップS61)。ここでは、処理の効率化のため、原画像を圧縮画像に変換して保存する。

【0129】次に、文書画像にラベリング処理を施し、矩形高さの最頻値をもとに大きな矩形を抽出し(ステップS62)、抽出された大きな矩形から表を囲む矩形(表矩形)を抽出して(ステップS63)、表矩形の中から管理情報を含むものを選択する(ステップS64)。ここでは、例えば、最も面積の大きな表矩形が選択される。

【0130】次に、選択された表矩形の内部から文字列を抽出し、文字列の外接矩形(文字列矩形)を求め、その座標をメモリに保存する(ステップS65)。次に、保存した文字列矩形から、横幅が小さい矩形や縦長矩形をノイズ矩形として除去し(ステップS66)、2つ以上の文字列矩形を統合する(ステップS67)。

【0131】ここまでの処理で、表内から抽出された文字列矩形が整理されるが、これらの文字列矩形は表罫線の一部を含んでいる場合もあり得る。そこで、文字列矩形の中の罫線部分を抽出し、その部分を境にして文字列

矩形を分割する（ステップS68）。

【0132】次に、管理情報に相当する文字列矩形を抽出するために、文字列矩形内の文字数を計算する（ステップS69）。ここで計算された文字数は、文字列矩形の属性としてステップS72の処理で用いられる。

【0133】ステップS68の処理により、表罫線で囲まれた欄毎の文字列矩形が抽出されるが、元の表の外形が矩形ではない場合には、表の外にある文字列矩形が残されている可能性がある。そこで、上罫線のチェックを行って、上側に表罫線がないような文字列矩形は表外の文字列矩形とみなし、それを除去する（ステップS70）。

【0134】次に、表内の文字列矩形を表矩形の左上座標に近い順に並び替える（ステップS71）。そして、文字列矩形の文字数が一定の条件を満たす場合に、その文字列矩形を管理情報として抽出して（ステップS72）、処理を終了する。条件を満たす文字列矩形が複数ある場合、表矩形の左上に近いものから順に管理情報の候補とする。

【0135】ここでは、処理結果の画像77において、管理情報の3つの候補C1、C2、C3が抽出されており、管理情報抽出装置のユーザインタフェース78は、これらを優先順位の高いものから順に出力して、ユーザに提示する。ユーザは、管理情報として適当な候補が提示された時点で、マウスのポインティング操作によりそれを選択する。もし、適当な候補が提示されないときは、ポインティング操作により他の矩形セルを明示的に指定することで、管理情報候補を修正する。

【0136】管理情報抽出装置は、ユーザが選択／修正した管理情報の位置を学習し、その位置情報と罫線構造を、ユーザ辞書73として辞書31内に格納しておく。これにより、管理情報抽出装置は、ユーザから直接指定された位置情報を、次回以降の処理で利用することができる。

【0137】また、図23の自動学習モードにおいては、管理情報抽出装置は、まず、複数のユーザ辞書73を参照しながら、図4の処理に従って入力画像71、入力画像72等のフォームを認識する。

【0138】ここで、表形式の画像71が入力され、大分類および詳細識別による照合の結果、いずれかのユーザ辞書73のフォームと一致すると判定した場合は、その処理結果のフォーム74においてあらかじめ指定されている位置の管理情報C1を出力し、それをユーザに提示する。ユーザが提示された管理情報C1を承認すれば、それをそのまま採用し、ユーザがそれを承認しなければ、他の位置の情報C2、C3等から適当なものを選択させる。

【0139】また、入力画像71がいずれのユーザ辞書73のフォームとも一致しない場合は、上述の表内管理情報抽出処理を行って、処理結果の画像75から管理情

報の候補C1、C2、C3等を抽出する。そして、ユーザインタフェース78は、これらを優先順位の高いものから順にユーザに提示し、ユーザは、それらのうち管理情報として適当なものを選択する。もし、適当な候補が提示されないときは、他の矩形セルを明示的に指定することで、管理情報候補を修正する。

【0140】管理情報抽出装置は、入力画像71においてユーザが選択／修正した管理情報の位置を学習し、その位置情報と罫線構造を、ユーザ辞書73として辞書31内に格納しておき、次回以降の処理で利用する。

【0141】また、表形式でない通常の文書画像72が入力された場合は、フォーム認識の結果、罫線構造がないと判定される。そこで、先願の罫線のない文書画像からのタイトル抽出処理を用いて、管理情報の複数の候補を抽出する。

【0142】図25は、この管理情報抽出処理のフローチャートである。処理が開始されると、管理情報抽出装置は、文書72を読み込み、文書画像としてメモリに格納する（ステップS81）。ここでは、処理の効率化のため、原画像を圧縮画像に変換して保存する。

【0143】次に、文書画像にラベリング処理を施し、その結果に基づいて文字列を抽出し、文字列矩形の座標をメモリに保存する（ステップS82）。次に、保存した文字列矩形から、横幅が小さい矩形や縦長矩形をノイズ矩形として除去し（ステップS83）、さらに文字列らしくない矩形を除いて、文書領域を決定する（ステップS84）。

【0144】次に、残った文字列矩形を縦方向（y座標）で並べ替え（ステップS85）、枠の画像を含む矩形（枠矩形）を抽出して、枠矩形内にある文字列矩形を枠付き矩形としてマークする（ステップS86）。また、下線の画像を含む矩形を抽出して、そのすぐ上にある文字列矩形を下線矩形としてマークする（ステップS87）。

【0145】次に、文字列矩形の文書内での位置、文字のサイズ、枠付き矩形や下線矩形に該当するかどうか等の性質に基づいて、タイトルらしさのポイント計算を行い、ポイントの高い1つ以上の文字列矩形をタイトル候補として抽出する（ステップS88）。そして、その結果を用いて文書の宛先と発信元情報を抽出する（ステップS89、S90）。こうして、タイトル、宛先、および発信元情報が管理情報の候補として抽出される。

【0146】ここでは、処理結果の画像76において、タイトルの3つの候補C4、C5、C6と、宛先および発信元情報が抽出されている。ユーザインタフェース78は、これらを優先順位の高いものから順に出力して、ユーザに提示する。ユーザは、管理情報として適当な候補が提示された時点で、マウスのポインティング操作によりそれを選択する。もし、適当な候補が提示されないときは、ポインティング操作により他の文字列矩形を明

示的に指定することで、管理情報候補を修正する。

【0147】次に、図26から図28までを参照しながら、こうして抽出された管理情報の利用形態について説明する。従来は、画像を管理する管理情報として、キーワードや文書名等の文字コードしか用いられていなかったが、本発明の管理情報抽出装置を備える電子ファイリングシステムでは、文字コード以外に文書画像の一部をインデックスとして格納する機能を備えておく。これにより、文字コードの信頼度が低い場合等に、画像による検索を行うことが可能になる。

【0148】このシステムは、管理情報を文字コードで保存するか画像コードで保存するかをユーザに選択させ、その選択結果に従って、いずれかのデータを管理情報として格納する。また、画像の検索時には、ユーザに管理情報の検索方法を選択させ、その結果に従って、文字コードまたは画像で検索する。また、単に文字コードまたは画像をバラバラと見るだけのブラウジングの機能も備えておく。

【0149】図26は、このような画像情報格納処理のフローチャートである。処理が開始されると、電子ファイリングシステムは、まず、文書画像を入力し（ステップS101）、例えば、図4の処理に従って管理情報の位置を算出して、管理情報の文字列を見つける（ステップS102）。そして、抽出した文字列に対する管理情報の保存方法をユーザに選択させる（ステップS103）。

【0150】保存方法としては、文字列を文字認識して文字コードに変換する文字認識モードと、文字認識を行わず文字列を画像で保存する画像モードがある。ユーザが文字認識モードを選択した場合は、文字認識を行い（ステップS104）、その認識結果の信頼度に応じて格納方法を選択する（ステップS105）。

【0151】文字認識の信頼度の算出方法としては、例えば、本出願人による先願の「文字認識方法および装置」（整理番号9604161）に記載された技術を用いる。この技術によれば、システムは、まず、認識結果の文字コードと入力された文字パターンとの距離値から確率パラメータを算出し、文字パターンと正解文字コードの集合から、確率パラメータを正読確率に変換する変換テーブルを生成する。次に、その変換テーブルを用いて、確率パラメータに対する正読確率を求め、その正読確率を認識結果の信頼度とする。

【0152】文字認識の信頼度が一定のしきい値より低い場合には、ユーザに画像で保存することを通知して、文字コードだけでなく、その文字列の画像も同時に管理情報として保存し（ステップS106）、処理を終了する。信頼度がそのしきい値以上の場合には、文字コードを管理情報として保存し（ステップS107）、処理を終了する。

【0153】また、ユーザが画像モードを選択した場合

は、文字列の画像を管理情報として保存し（ステップS108）、処理を終了する。ステップS103において、他の保存方法として、文字コードと画像コードの両方を保存するモードを設けることも、もちろん可能である。また、ステップS105における信頼度として、認識結果の文字コードと入力された文字パターンとの距離値の情報をを用い、距離値が小さいほど信頼度が高いものとして判定を行ってもよい。

【0154】図27は、管理情報を格納する格納テーブルの例を示している。図27の管理情報格納テーブルでは、文字コード格納領域、画像格納領域、および文字コードで保存するか画像コードで保存するかの区別を表す種別フラグ領域（flag）が設けられている。

【0155】例えば、種別フラグの値が“0”のときは、文字コードのみが保存されていることを表し、その値が“1”のときは、画像コードのみが保存されていることを表し、その値が“2”のときは、文字コードと画像コードの両方が保存されていることを表す。

【0156】また、図28は、このような管理情報を検索する管理情報検索処理のフローチャートである。処理が開始されると、電子ファイリングシステムは、まず、管理情報の検索方法をユーザに選択させる（ステップS111）。検索方法としては、文字コードで検索するモード、画像で検索するモード、および文字コードと画像のリストを表示して、ユーザにブラウジングさせるモードの3つのモードがある。

【0157】ユーザが文字コード検索を選択すれば、文字コードにより管理情報を検索し（ステップS112）、画像検索を選択すれば、画像により管理情報を検索し（ステップS113）、ブラウジングを選択すれば、管理情報格納テーブルに格納されている文字コードと画像のリストを表示する（ステップS114）。そして、処理を終了する。

【0158】ステップS113において画像で検索する場合は、ユーザに特定の画像のファイルを指定させるか、適当な画像を選んで表示する。そして、その中の特定の矩形部分を検索キーとしてユーザに指定させ、ユーザが指定した部分画像と管理情報格納テーブルに保存してある画像とを照合する。これらの画像同士の照合には、例えば、「画像理解のためのデジタル画像処理 [I]」（鳥脇純一郎著、昭晃堂）に記載されているような公知のテンプレートマッチングを用いる。

【0159】テンプレートマッチングでは、指定された部分画像をモデル（テンプレート）として用い、それと各管理情報格納テーブル内の画像との類似度を計算する。そして、類似度が最大となる管理情報、あるいは類似度が一定値以上の管理情報を求める。そして、得られた管理情報に対応する文書画像を検索結果として表示する。

【0160】このような電子ファイリングシステムによ

れば、管理情報の文字列を文字コードで格納／検索するだけでなく、画像そのもので格納／検索することもできるので、テキスト付き文字やデザインフォント、ロゴ等の正確に認識しにくい文字でも、管理情報として扱うことができる。

【0161】ところで、図17のステップS15、S16においては、表形式文書のフォーム（フォーマット構造）を識別するために、CROSS RATIO のDPマッチングを用いたが、詳細識別は他の任意の方法により行うこともできる。

【0162】他の公知のフォーム自動識別法においては、まず、既知の表形式文書のフォームの特徴量を、モデルとして辞書31に登録しておく。そして、未知の表形式文書の画像が入力されたときに、その画像から特徴量を算出し、辞書のモデルとマッチングを行って、最も類似するモデルを特定する。

【0163】モデルマッチングの1つの方法としては、表全体を正規化後、各矩形セルの中心点の位置を算出し、ほぼ同じ位置に中心点を持つモデルに投票を行い、最大投票数を得たモデルを最適モデルとする方法がある。ここで、表の正規化とは、表の縦横比が1対1になるように画像全体を変換する等の調整処理を意味する。

【0164】また、もう1つの方法として、連合グラフを用いたマッチングがある。この方法では、まず、罫線を抽出してから表全体を正規化し、入力された未知文書と各モデルとの間で、ほぼ同じ位置かつ同じ長さの罫線同士の組み合わせを求める。そして、図29に示すように、それらの各組み合わせを表すノードを平面上に配置し、一定の幾何的制約を満たすノード間をパスで結んで、連合グラフを生成する。

【0165】ここで、幾何的制約とは、比較する未知文書とモデルの間で罫線の順序関係を保存するという制約条件や、一方の表の1つの罫線に他方の表の複数の罫線が重複して対応することを禁止する制約条件等を指す。図29の4つのノードから成る連合グラフにおいては、未知文書の罫線a1、a2、a3、a4は、それぞれ、モデルの罫線b1、b2、b3、b4に対応している。

【0166】また、連合グラフの一部の部分グラフにおいて、どのノードも他のすべてのノードとパスで結ばれているとき、その部分グラフはクリークと呼ばれる。図29の連合グラフは、それ自身がクリークとなっている。連合グラフからノード数が最大の最大クリークを求めることで、未知文書とモデルの類似度が求められ、最大類似度を持つモデルが最適モデルとして抽出される。

【0167】これらのモデルマッチングでは、いずれも、未知の入力文書を正規化後、モデルと重ね合わせて特徴量を比較することが重要なステップとなっている。しかし、表の外形形状の抽出精度が劣化したり、行の追加等による若干のフォーム変更が行われたりした場合、特徴量全体が影響を受けるため、識別が不安定になる。

特に、矩形セルの中心位置を用いる前者の方法においてその影響は大きい。

【0168】連合グラフを用いる後者の方法においては、ノードを生成するための条件を緩くすることで対処できるが、グラフの規模が大きくなるため、とりわけ最大クリークを求める処理に時間がかかってしまう。

【0169】そこで、以下の実施形態では、罫線のかすれやノイズによる罫線抽出の誤り、若干のフォーム変更等に対して、ロバスト性（強靱さ）が優れ、しかも高速なマッチング方法について述べる。ここで、ロバスト性が優れているとは、マッチング結果が誤りや変更の影響を受けにくいことを意味する。

【0170】このマッチング方法では、まず、罫線の表全体に対する大きさ、位置を特徴量として、未知文書と各モデルとの間で罫線と罫線の対応の可能性を調べ、対応する罫線の組み合わせを求める。ここでは、可能性条件を広くとることで、1本の罫線に対し複数の罫線が対応可能とする。このように重複を許すことで、表の正しい外形形状の抽出に失敗した場合でも、その失敗がある程度以下であれば、罫線間の正しい対応関係を見落とすことがなくなる。

【0171】次に、得られた対応関係の集合の中から、互いに両立するものを1つのグループに集め、各対応関係をいくつかのグループに振り分ける。この際、特徴量としては、罫線と罫線の間の大きさの関係や位置関係等を用いる。罫線間の相対的な関係を特徴量として用いることで、かすれやノイズが特徴量全体にわたって大きな影響を及ぼすことがなくなる。

【0172】さらに、グループ分けの際に、両立の性質が推移的になるように、両立の条件を強くすることで、両立の可否を調べる処理回数が上述の連合グラフをつくる場合に比べて大幅に少なくなる。また、各グループに含まれる対応関係は、平面上の1本のパスで表現することができるので、その数を数える時間もほとんどかからない。

【0173】ここで、両立が推移的とは、例えば、対応Aと対応Bが両立し、対応Bと対応Cが両立するとき、必ず対応Aと対応Cが両立することを意味する。この場合には、改めて対応Aと対応Cの両立性をチェックする必要がないので、処理が高速化される。

【0174】そして、最後に、得られたグループ間の矛盾のない組み合わせの中で、含まれる対応関係の数が最大となる組み合わせを探索する。これにより、ある表に行を1行だけ追加する等の若干の変更がある場合でも、他の大部分の罫線が正しく対応していれば、そのモデルを抽出することが可能になる。

【0175】図30は、このようなマッチング方法を用いたフォーム識別処理のフローチャートである。この処理は、図4のステップS11、S12、S15、S16の処理に対応し、特に、入力画像の詳細識別に関する。

処理が開始されると、管理情報抽出装置は、まず、画像を入力し（ステップS121）、入力された画像から野線を抽出する（ステップS122）。

【0176】そして、その野線を囲む矩形（野線矩形）の左上頂点の座標に関し、横野線はy座標の小さい順に、縦野線はx座標の小さい順に、各野線を並べ直す（ステップS123）。ここで、横野線のy座標が同じ場合は、そのx座標の小さい順にソートし、縦野線のx座標が同じ場合は、そのy座標の小さい順にソートする。

【0177】次に、縦横各野線について概略情報を抽出する（ステップS124）。概略情報とは、表全体に対する野線の長さおよび位置の相対的な値であり、3つの整数の組で表される。また、縦横各方向における2本の野線のすべての組み合わせを考え、各組み合わせに関する詳細情報を抽出する（ステップS125）。詳細情報は、2本の野線間の長さおよび位置の相対的な関係を表す。

【0178】入力画像と照合されるモデルの概略情報および詳細情報は、あらかじめ抽出されてレイアウト辞書31に格納されている。そこで、次に、入力画像の概略情報および詳細情報と、モデルの概略情報および詳細情報とを照合し、モデルマッチングを行う（ステップS126）。そして、最適なモデルを識別結果として出力し（ステップS127）、処理を終了する。

【0179】次に、図31から図41までを参照しながら、ステップS124、S125、S126、およびS127の処理について詳細に説明する。ステップS124においては、概略情報を求めるための前処理として、表の基準幅W、基準高H、基準x座標x0、および基準

y座標y0を求める。まず、横野線に関してそれらの最大長を求め、その長さに比べてあるしきい値以上の割合（例えば0.8）の長さを持つ横野線のうち、順位が最初のものと最後のものとを求め、それらを外郭基準横野線とする。

【0180】縦野線に関しても最大長を求め、横野線の場合と同様にして、2本の外郭基準縦野線を得る。そこで、得られた4本の外郭基準野線の外接矩形を考え、その幅を基準幅W、その高さを基準高H、その左上頂点を基準点とし、その座標を基準座標（x0, y0）とする。

【0181】例えば、図31のような表形式文書においては、横野線81、82が外郭基準横野線として抽出され、縦野線83、84が外郭基準縦野線として抽出され、これらの外郭基準野線の外接矩形の幅が基準幅Wとなり、その高さが基準高Hとなる。また、その外接矩形の左上頂点85の座標が基準座標（x0, y0）となる。

【0182】ここで、最大長から算出される一定長さ以上の野線の中から外郭基準野線を選ぶことで、例えば、横野線86、87のような短い野線が外郭基準野線の候補から除外される。

【0183】あるいはまた、これらの基準幅W、基準高H、基準座標（x0, y0）を、次のようにして求めることも可能である。まず、基準座標の候補となる座標値vmaxx、vminx、vmaxy、vminy、hmaxx、hminx、hmaxy、hminyを、次式のように定義する。

【0184】

$$\begin{aligned} \text{vmaxx} &= (\text{縦野線矩形の右下頂点のx座標の最大値}) \\ \text{vminx} &= (\text{縦野線矩形の左上頂点のx座標の最小値}) \\ \text{vmaxy} &= (\text{縦野線矩形の右下頂点のy座標の最大値}) \\ \text{vminy} &= (\text{縦野線矩形の左上頂点のy座標の最小値}) \\ \text{hmaxx} &= (\text{横野線矩形の右下頂点のx座標の最大値}) \\ \text{hminx} &= (\text{横野線矩形の左上頂点のx座標の最小値}) \\ \text{hmaxy} &= (\text{横野線矩形の右下頂点のy座標の最大値}) \\ \text{hminy} &= (\text{横野線矩形の左上頂点のy座標の最小値}) \end{aligned} \quad (10)$$

次に、これらの座標値から、次式により基準幅と基準高の候補を求める。

$$\begin{aligned} W1 &= \text{vmaxx} - \text{vminx} \\ W2 &= \text{hmaxx} - \text{hminx} \\ H1 &= \text{hmaxy} - \text{hminy} \\ H2 &= \text{vmaxy} - \text{vminy} \end{aligned} \quad (11)$$

そして、基準幅Wを、

$$W = \max \{W1, W2\} \quad (12)$$

により求め、 $W=W1$ のとき、 $x0 = \text{vminx}$ とし、 $W=W2$ のとき、 $x0 = \text{hminx}$ とする。

$$H = \min \{H1, H2\} \quad (13)$$

により求め、 $H=H1$ のとき、 $y0 = \text{hminy}$ とし、 $H=H2$ のとき、 $y0 = \text{vminy}$ とする。

【0186】また、基準高Hを、

【0187】こうして、基準幅W、基準高H、基準座標(x0, y0)が求められる。ただし、この方法は、上述の方法に比べてノイズ等の影響を受けやすく、ロバスト性が比較的弱くなる。

【0188】ここでは、4本の外郭基準罫線の外接矩形の左上頂点を基準点として選んだが、それ以外に、左下頂点、右上頂点、右下頂点等、外接矩形の周辺上の任意の点を基準点として選ぶことができる。その場合でも、

$$\begin{aligned} \text{length1} &= [(l1/W) \times 100] \text{ の整数部分} \\ \text{twist} &= [((x1-x0) / W) \times 100] \text{ の整数部分} \\ \text{position} &= [((y1-y0) / H) \times 100] \text{ の整数部分} \end{aligned} \quad (14)$$

また、縦罫線の場合、図33に示すように、罫線矩形92の長さl1と中心座標(x1, y1)から、次式によ

$$\begin{aligned} \text{length1} &= [(l1/H) \times 100] \text{ の整数部分} \\ \text{twist} &= [((y1-y0) / H) \times 100] \text{ の整数部分} \\ \text{position} &= [((x1-x0) / W) \times 100] \text{ の整数部分} \end{aligned} \quad (15)$$

得られた特徴量のうち、length1は表のサイズに対する罫線の長さの相対的な割合を表し、twistとpositionは、表の基準点に対する罫線の相対的な位置を表している。

【0192】次に、ステップS125においては、2本の罫線の相対的な関係を表す詳細情報を求める。ここでは、一方の罫線矩形の長さを1としたときの、もう一方の罫線矩形の長さlength2、それぞれの罫線矩形の中心間のx方向のずれの長さdiff er、および中心間のy方向のずれの長さheightの3つの値を、詳細情報として求めることにする。

$$\begin{aligned} dw &= x2 - x1 \\ dh &= y2 - y1 \end{aligned}$$

この定義によれば、罫線矩形94の中心が罫線矩形93の中心より右にあれば、dwは正となり、罫線矩形94の中心が罫線矩形93の中心より左にあれば、dwは負となる。同様に、罫線矩形94の中心が罫線矩形93の中心より下にあれば、dhは正となり、罫線矩形94の

$$\begin{aligned} \text{length2} &= l2 / l1 \\ \text{diff er} &= dw / l1 \\ \text{height} &= dh / l1 \end{aligned}$$

また、縦罫線についても同様に、2本の罫線のすべての組み合わせを抽出する。そして、各組み合わせにおいて、図35に示すように、ソート順位の早い方の罫線矩形95の長さをl1、その中心座標を(x1, y1)、ソート順位の遅い方の罫線矩形96の長さをl2、その

$$\begin{aligned} \text{length2} &= l2 / l1 \\ \text{diff er} &= dh / l1 \\ \text{height} &= dw / l1 \end{aligned}$$

(17-2)式においては、diff erとheightの定義が、(17-1)式と逆になっている。次に、ステップS126において、横罫線同士と縦罫線同士の2回に分けて、入力画像の概略情報および詳細情報を、各モデルの概略情報および詳細情報と照合し、フォーム

以下の処理は基本的に同様である。

【0189】次に、得られた表の大きさと基準座標をもとにして、各罫線矩形の長さと中心位置の情報から、3つの特徴量(概略情報)length1、twist、positionを求める。横罫線の場合、図32に示すように、罫線矩形91の長さl1と中心座標(x1, y1)から、次式によりこれらの特徴量を算出する。

【0190】

length1、twist、positionよりこれらの特徴量を算出する。

【0191】

【0193】まず、2本の横罫線のすべての組み合わせを抽出する。そして、各組み合わせにおいて、図34に示すように、一方の罫線矩形(ソート順位の早い方)93の長さをl1、その中心座標を(x1, y1)、もう一方の罫線矩形(ソート順位の遅い方)94の長さをl2、その中心座標を(x2, y2)とする。このとき、罫線矩形93の中心を基準として、これらの罫線矩形の中心間のx方向のずれdwとy方向のずれをdhを次式により定義する。

【0194】

$$(16)$$

中心が罫線矩形93の中心より上にあれば、dhは負となる。

【0195】次に、上述の3つの特徴量length2、diff er、heightを、次式により算出する。

$$(17-1)$$

中心座標を(x2, y2)とする。そして、(16)式によりdwとdhを求め、次式により、詳細情報length2、diff er、heightを算出する。

【0196】

$$(17-2)$$

の類似度を計算する。

【0197】図36は、このようなモデルマッチング処理のフローチャートである。処理が開始されると、管理情報抽出装置は、まず、未知文書の入力画像の横罫線の数をp本、モデルの横罫線の数をm本として、図37に

示すような $p \times m$ の表を作成する (ステップ S 131)。

【0198】ここでは、 $p=12$ 、 $m=15$ であり、表の行番号および列番号は、ともに0から始まっている。この表の i 行 j 列の要素 (項目) は、入力画像の i 番目の野線とモデルの j 番目の野線の対応関係を表すデータである。以下では、このような表をマッチングテーブルと呼ぶことにする。

【0199】次に、概略情報に基づいて、入力画像の i 番目の横野線 $IP(i)$ が、モデルの j 番目の横野線 $MO(j)$ に対応するかどうかの可能性を判断し、対応する可能性があれば、マッチングテーブルの i 行 j 列の要素にノードを配置する (ステップ S 132)。これにより、横野線 $IP(i)$ と横野線 $MO(j)$ の組み合わせがマッチングテーブル上に記述される。このときの対応

$$\begin{aligned} |ipl - mol| &< \beta \\ |ipt - mot| &< \beta \\ |ipp - mop| &< \alpha \end{aligned}$$

これらの野線の数に依存したパラメータ α 、 β は正の整数で、野線数が少ないほど大きな値をとり、逆に野線数が多いほど小さな値をとるように定められる。このとき、(18) 式の条件は、表内の野線の密度が疎であればマッチングの探索範囲を広げ、密度が密であれば探索範囲を狭めるという効果を持つ。 α 、 β は、例えば、図38に示すように、横、縦の野線数に対して単調に減少する関数として定義できる。

【0203】このように、概略情報が互いに類似している野線同士の対応関係をノードで表現することで、入力画像とモデルの間の、表の外郭部分に対する相対的な特徴の類似点が抽出される。

【0204】次に、詳細情報に基づいて、配置されたノードの中で、互いに特定の関係を満たすもの同士、すなわち両立するもの同士の組み合わせを探索する (ステップ S 133)。そして、それらのノードを同じグループに属するノードとみなして、パスで結ぶ。

【0205】ここで、 i 行 j 列のノード $n(i, j)$ と k 行 l 列のノード $n(k, l)$ が特定の関係を満たすとは、入力画像の i 番目の野線と k 番目の野線の関係が、モデルの j 番目の野線と l 番目の野線の関係に比例していることを指す。言い換えれば、入力画像の i 番目の野線とモデルの j 番目の野線をぴったり重ね合わせたと、入力画像の k 番目の野線とモデルの l 番目の野線がぴったり重なり合う状況を指す。

【0206】このようなノード同士を1本のパスで結ぶことで、ノードがいくつかのグループに分類される。ノードの数が多いグループほど、入力文書とモデルが類似していることを強く表していると考えられるので、このようなグループを対象とすることで、モデルマッチングにおける類似度の計算を効率よく行うことができる。

【0207】また、あるノードと両立するノードを探索

可能性の条件は十分緩くおき、一つの野線に対して、複数の野線が重複して対応することを許すものとする。

【0200】ここでは、野線 $IP(i)$ の概略情報 ($length, twist, position$) をそれぞれ (ipl, ipt, ipp) とおき、野線 $MO(j)$ の概略情報をそれぞれ (mol, mot, mop) とおいて、各値の差が一定値より小さいとき、野線 $IP(i)$ と野線 $MO(j)$ が対応する可能性があるとみなす。

【0201】可能性を表す具体的な条件は、表内の横野線の数に依存したパラメータ α 、および縦野線の数に依存したパラメータ β をしきい値として、次式により与えられる。

$$\begin{aligned} \text{【0202】} \\ (18) \end{aligned}$$

する際、処理の効率を高めるため、常に前者のノードの右下の領域に位置するノードを探索対象とする。これにより、図29に示したようなクリークが生成されなくなるので、ノード数の多いパスを高速に求めることが可能になる。パス生成の具体的な処理については後述する。

【0208】次に、得られたパスの集合の中で、互いに矛盾しないパスの組み合わせを求め、いくつかの組み合わせのうち、含まれるノード数が最大のものを探索する (ステップ S 134)。そして、そのパスの組み合わせを最適パス集合とする。2つのパスが互いに矛盾しないとは、それぞれのパスの中のノードに対応する野線の集合の範囲が、互いに重なり合わないということである。

【0209】図37のマッチングテーブルにおいて、2つの野線集合の範囲が重なり合う場合は、図39に示すように、入力画像とモデルの間で対応する野線の順序関係が逆転する場合と、図40に示すように、野線と野線の対応が重複する場合の2通りある。

【0210】図39においては、実線で表されたグループに属するモデル側の野線の範囲は0番目から9番目となっており、破線で表されたグループに属するモデル側の野線の範囲は7番目から8番目となっている。したがって、2つの野線集合の範囲が互いに重なっている。同様に、図40においても、実線で表されたグループと破線で表されたグループの野線集合の範囲が、モデル側において互いに重なっている。

【0211】一方、矛盾したパスの組み合わせを含まない最適パス集合においては、図41に示すように、入力画像とモデルの両方の側において、各野線集合の範囲は重なり合うことがない。このような最適パス集合に含まれる各ノードによって表される野線間の対応関係が、最適な対応関係である。

【0212】次に、入力画像の横野線数を p_h 、モデル

の横野線数を m_h 、横野線の最適パス集合に含まれるノード数を max_h として、入力画像とモデルの横野線の

$$SH = max_h / p_h + max_h / m_h$$

この類似度SHは、入力画像の横野線のうち、最適パス集合に対応するものの割合と、モデルの横野線のうち、最適パス集合に対応するものの割合の和を表し、一般に、入力画像とモデルの横野線の特徴が類似しているほど大きな値をとる。

【0214】管理情報抽出装置は、以上説明したステッ

$$SV = max_v / p_v + max_v / m_v$$

そして、最終的に、SHとSVから、次式により入力画像とモデルの野線の類似度Sを算出し、モデルマッチン

$$S = SH + SV$$

例えば、図4のステップS14の大分類により得られたすべての表の候補をモデルとして、このようなモデルマッチングを行うことで、各モデルと入力画像の類似度が算出される。ステップS127においては、これらのモデルのうち最も類似度の高いものを、最適モデルとして出力する。これにより、入力画像に対応する辞書フォームが特定される。

【0217】次に、図42から図48までを参照しながら、図36のノード配置処理、パス生成処理、および最適パス集合決定処理について、さらに詳しく説明する。図42は、図36のステップS132におけるノード配置処理のフローチャートである。ここでは、入力画像のi番目の野線の概略情報length1、twist、positionを、それぞれipl(i)、ipt(i)、ipp(i)とおき、モデルのj番目の野線の概略情報を、それぞれmol(j)、mot(j)、mop(j)とおく。

【0218】また、マッチングテーブルのi行j列の要素を表すデータをsign(i, j)とおく。sign(i, j) = 0のときは、対応する要素にノードは配置されず、sign(i, j) = 1のときは、その要素にノードが配置される。

【0219】処理が開始されると、管理情報抽出装置は、まず、条件|ipp(i) - mop(j)| < αが成り立つかどうかを判定する(ステップS141)。この条件が成り立たなければ、sign(i, j) = 0において(ステップS142)、処理を終了する。

【0220】ステップS141の条件が成り立てば、次に、条件|ipt(i) - mot(j)| < βが成り立つかどうかを判定する(ステップS143)。この条件が成り立たなければ、sign(i, j) = 0において(ステップS144)、処理を終了する。

【0221】ステップS143の条件が成り立てば、次に、条件|ipl(i) - mol(j)| < βが成り立つかどうかを判定する(ステップS145)。この条件が成り立たなければ、sign(i, j) = 0において(ステップS146)、処理を終了する。そして、ステ

類似度SHを次式により求める(ステップS135)。
【0213】

$$(19)$$

プS131からステップS135までと同様の処理を縦野線についても行う。そして、入力画像の縦野線数をp_v、モデルの縦野線数をm_v、縦野線の最適パス集合に含まれるノード数をmax_vとして、入力画像とモデルの縦野線の類似度SVを次式により求める。

【0215】

$$(20)$$

グを終了する。

【0216】

$$(21)$$

ップS145の条件が成り立てば、sign(i, j) = 1において、i行j列にノードを配置し(ステップS147)、処理を終了する。

【0222】このような処理を、マッチングテーブルのすべての位置(i, j)について行うことで、概略情報が類似した2つの野線に対応する位置に、それらの間の対応関係を表すノードが配置される。

【0223】また、図43、44は、図36のステップS133におけるパス生成処理のフローチャートである。処理が開始されると、管理情報抽出装置は、まず、初期化処理を行う(図43、ステップS151)。ここでは、マッチングテーブル上でノードが配置された要素の位置(i, j)を、ノード列としてメモリ内の記憶部に格納する。記憶部内でのノードの並べ方は、行番号iの小さい順とし、同じiに対しては列番号jの小さい順とする。また、このノード列の各ノードには、パスで結ばれたかどうかを示すフラグが付加される。

【0224】例えば、図37のマッチングテーブルに対応する記憶部のノード列は、図45に示すようになる。図45の記憶部においては、マッチングテーブル上の各ノードの位置(0, 0)、(1, 0)、(1, 1)、(2, 0)、・・・、(11, 14)が順に格納され、それらのフラグの値は1に初期化されている。フラグの値が1のとき、対応するノードはまだパスで結ばれていないことを表す。

【0225】次に、記憶部の先頭のデータにアクセスし(ステップS152)、そのアクセスポイントからiとjを読み出して、その位置に対応するマッチングテーブル上の要素をマークする(ステップS153)。そして、マークした要素のノードを基準ノードとして、その要素のsignを0にし、対応する記憶部のフラグを0にする(ステップS154)。

【0226】次に、制御変数countの値を0とおき(ステップS155)、マークした要素がマッチングテーブルの最終列に対応しているか、または、countの値があらかじめ決められた定数hに達したかを調べる(ステップS156)。これらの条件がいずれも満たさ

れなければ、次に、マーク位置を1列分右に進め（ステップS157）、マーク位置が最終行に対応しているかどうかを調べる（ステップS158）。

【0227】ここで、マーク位置が最終行であれば、countの値に1を加算して（ステップS159）、次の列の要素を調べるために、ステップS156以降の処理を繰り返す。マーク位置が最終行でなければ、次に、そのマーク位置を1行分下に進め（ステップS160）、ここでマークした要素のsignが0か1かを判定する（ステップS161）。

【0228】その値が0であれば、マーク位置にノードは配置されていないので、同じ列内の他の要素を調べるために、ステップS158以降の処理を繰り返す。signが1であれば、マーク位置にノードが配置されているので、そのノードが基準ノードとパスで結べるかどうかを判定する（ステップS162）。2つのノードがパスで結べるかどうかは、これらのノードに対応する野線

$$\begin{aligned} |l2/l1 - l2'/l1'| &< \varepsilon 1 \\ |dw/l1 - dw'/l1'| &< \varepsilon 2 \\ |dh/l1 - dh'/l1'| &< \varepsilon 3 \end{aligned}$$

しきい値 $\varepsilon 1$ 、 $\varepsilon 2$ 、 $\varepsilon 3$ を十分小さくすることで、

(22)式は、野線101と野線102から成る図形が、野線103と野線104から成る図形と相似であることを表すようになる。これらの野線図形が相似であれば、野線101と野線103を対応させたとき、同時に、野線102と野線104が対応する可能性が高いので、2つのノードは両立するとみなされる。

【0233】このように、パスを引く条件として相似条件を用いることで、ノード間の両立性を判定する回数を削減することができる。例えば、図37のマッチングテーブルにおいて、ノード97を基準ノードとした場合、ノード97とノード98が両立し、かつ、ノード97とノード99が両立すれば、ノード98とノード99も両立するとみなしてよい。

【0234】そこで、ノード99が基準ノード97とパスで結べると判定したとき、そのノード99はすでに基準ノード97とパスで結ばれているノード98ともパスで結べるものと判断する。

【0235】マーク位置のノードが基準ノードとパスで結べない場合は、同じ列内の他のノードを調べるために、ステップS158以降の処理を繰り返す。また、それらがパスで結べる場合は、そのマーク位置のノードに対応する記憶部内のフラグを0に書き換える（ステップS163）。これにより、そのノードが、基準ノードまたはすでにパス上に存在する1つ前のノードと結ばれたことが記録される。そして、次の列のノードを調べるために、ステップS156以降の処理を繰り返す。

【0236】ステップS156以降の処理では、マーク位置を1列進めた後、直ちに1行進めることで、常に、前のマーク位置の右下にある要素を探索対象とする。こ

間の詳細情報length2、differ、heightを用いて判定される。

【0229】例えば、図46に示すように、入力画像において、基準ノードに対応する野線101と判定対象のノードに対応する野線102との間の相対的な関係を表す詳細情報を、 $length2 = l2/l1$ 、 $differ = dw/l1$ 、 $height = dh/l1$ とおく。

【0230】また、モデルにおいて、基準ノードに対応する野線103と、判定対象のノードに対応する野線104との間の相対的な関係を表す詳細情報を、 $length2 = l2'/l1'$ 、 $differ = dw'/l1'$ 、 $height = dh'/l1'$ とおく。

【0231】このとき、経験的に決められるしきい値 $\varepsilon 1$ 、 $\varepsilon 2$ 、 $\varepsilon 3$ に対して次式が成り立てば、基準ノードと判定対象のノードが両立し、これらの間をパスで結べるものとする。

【0232】

(22)

れを繰り返すことで、マッチングテーブルの右下に向かって、徐々にパスが伸張されていく。

【0237】そして、ステップS156の条件が満たされると、次に、基準ノードから伸びたパスのヒット数が2以上かどうかを調べる（図44、ステップS164）。ここで、ヒット数とは、パスの構成ノードの数を意味する。構成ノードの数が2以上であれば、そのパスを正式に登録し、その構成ノードの情報を記録する（ステップS165）。パスの構成ノードの数が1の場合は、基準ノードからいずれの他のノードにもパスが引かれなかったことを意味するので、パス登録は行わない。

【0238】次に、記憶部内にアクセスされていないデータが残っているかどうかを調べる（ステップS166）。データが残っていれば、記憶部のアクセスポイントの一つ進めて（ステップS167）、その位置のフラグの値を調べる（ステップS168）。そして、フラグが0であれば、その位置のノードはすでにパス上に加えられているので、ステップS166以降の処理を繰り返して次のデータを調べる。

【0239】また、フラグが1であれば、その位置のノードはまだパスに加えられていないので、ステップS153以降の処理を繰り返す。これにより、そのノードを新たな基準ノードとする新しいパスが生成される。そして、ステップS166において、記憶部のアクセスポイントが最後尾に達すれば、処理を終了する。

【0240】また、図47は、図36のステップS134における最適パス集合決定処理のフローチャートである。この処理においては、p行m列の横野線または縦野線のマッチングテーブルを対象とし、暫定的な最適パス集合のノード数を表す配列score(i) (i=0,

1, 2, . . . , m) と、行番号を表す配列 $r_{irek i(i)}$ ($i=0, 1, 2, . . . , m$) を用いる。

【0241】処理が開始されると、管理情報抽出装置は、まず、最適パス集合のノード数の初期値を表す $score(m)$ を0とおき、行番号の初期値を表す $r_{irek i(m)}$ を $p-1$ とおく (ステップS171)。

【0242】次に、列番号を表す変数 i を $m-1$ とおく (ステップS172)、登録されたパスの中で、始点となる左上のノードの列番号が i であるようなパスの集合を $Path(i)$ とする (ステップS173)。また、 $score(i) = score(i+1)$ 、 $r_{irek i(i)} = r_{irek i(i+1)}$ とおく (ステップS174)。この $score(i)$ は、 i 番目の列から最終列 ($m-1$ 番目の列) までの範囲における暫定的な最適パス集合のノード数を表す。

【0243】次に、集合 $Path(i)$ の1つのパスを取り出し、そのノードの情報を用いて $score(i)$ を更新する (ステップS175)。そして、集合 $Path(i)$ にパスが残っているかどうかを調べ (ステップS176)、残っていれば、次のパスを取り出して $score(i)$ の計算を繰り返す。

【0244】集合 $Path(i)$ のすべてのパスに関する計算が終了すると、次に、 i が0に達したかどうかを判定し (ステップS177)、 i が1以上であれば、 $i = i-1$ とおいて (ステップS178)、ステップS173以降の処理を繰り返す。そして、 i が0に達したとき、得られた $score(0)$ の値を最終的な最適パス集合のノード数として (ステップS179)、処理を終了する。

【0245】横野線のマッチングテーブルから得られた $score(0)$ の値は、類似度の計算において、(19) 式の max_h として用いられ、縦野線のマッチングテーブルから得られた $score(0)$ の値は、(20) 式の max_v として用いられる。

【0246】次に、図48を参照しながら、図47のステップS175のノード数更新処理について説明する。ノード数更新処理が開始されると、管理情報抽出装置は、まず、集合 $Path(i)$ の1つのパスを取り出し、そのパスの始点の行番号を sg とし、終点となる右下のノードの列番号、行番号をそれぞれ er 、 eg とし、パスに含まれるノードの数を $hits$ とする (ステップS181)。

【0247】例えば、図37のマッチングテーブルにおいて、 $i=11$ のとき、 $Path(11)$ には右下の領域のパス $p1$ 、 $p2$ が含まれる。このうち、パス $p1$ については、 $sg=8$ 、 $er=14$ 、 $eg=11$ となり、パス $p2$ については、 $sg=6$ 、 $er=12$ 、 $eg=7$ となる。

【0248】次に、列番号を表す変数 j を $er+1$ とおく (ステップS182)、 eg と $r_{irek i(j)}$ の

値を比較する (ステップS183)。ここで、 eg が $r_{irek i(j)}$ より大きければ、次に、 $score(j) + hits > score(i)$ 、または、 $score(j) + hits = score(i)$ かつ $eg < r_{irek i(i)}$ が成り立つかどうかを判定する (ステップS184)。

【0249】そして、これらの条件のいずれかが成り立てば、 $score(i) = score(j) + hits$ 、 $r_{irek i(i)} = eg$ とおいて (ステップS185)、処理を終了する。

【0250】また、ステップS183において eg が $r_{irek i(j)}$ 以下のとき、および、ステップS184においていずれの条件も成り立たないときは、 $j = j+1$ とおいて (ステップS186)、 j と m を比較する (ステップS187)。ここで、 j が m 以下であれば、ステップS183以降の処理を繰り返し、 j が m を越えようと、処理を終了する。

【0251】こうして、1つ前の処理で暫定的に決められた最適パス集合にパスを1つ加えてできる集合から、新たな最適パス集合が抽出され、そのノード数が $score(i)$ に記録される。このような処理を、 $Path(i)$ のすべてのパスについて繰り返すことで、 i 番目の列から最終列までの範囲における最適パス集合のノード数が得られる。

【0252】例えば、図37において、11番目の列から最終列までの範囲における互いに矛盾しないパスの組み合わせとしては、パス $p1$ のみと、パス $p2$ および $p3$ の組み合わせの2つが考えられる。これらの組み合わせのノード数はともに4なので、 $score(11) = 4$ となる。

【0253】以上説明したフォーム識別処理は、管理情報抽出装置だけでなく、文書認識装置、図面読み取り装置等の任意の画像認識装置にも適用でき、任意の画像中の野線構造を識別する際に有効である。

【0254】本実施形態のフォーム識別処理によれば、野線間の相対的な関係の特徴量として用いているので、入力された表形式文書から野線構造を抽出し、登録された表形式文書のフォームとマッチングを行う際に、かすれやノイズ等の影響で野線がうまく抽出されない部分があっても、安定して正しい識別が行われる。とりわけ、ノイズの影響で野線抽出が不安定になりやすい外郭野線抽出の精度劣化に対しても、ノード配置のための条件を広くとることで、ロバスト性を保つことができる。

【0255】また、1つ以上のパスの組み合わせである最適パス集合を求めることで、1行追加、削除などの若干のフォーム変更に対しても、安定して正しい識別が行われる。さらに、2つのノードに関する両立の条件を推移的とすることで、両立性のチェック回数が削減され、高速な識別処理が可能になる。

【0256】

【発明の効果】本発明によれば、表形式文書等の画像のフォームと管理情報の位置が自動的に学習され、辞書に記憶されるので、それらの情報を用いて、任意の入力画像における管理情報の位置を高精度に算出することができる。

【0257】特に、画像情報の変動に強い特徴量を用いているので、かすれや潰れのある文書画像でも管理情報の抽出が可能になる。また、フォームの学習／照合を大分類と詳細識別の2段階で候補を絞りながら処理を実行し、詳細識別では表の外側の形状の特徴を用いて1次元

マッチングで識別するので、高速に管理情報が抽出される。

【0258】さらに、管理情報を文字コードで格納／検索するだけでなく、画像そのもので格納／検索することもできるので、テキスト付き文字等の正確に文字認識しにくい文字でも、管理情報として扱うことが可能になる。

【図面の簡単な説明】

【図1】本発明の管理情報抽出装置の原理図である。

【図2】管理情報抽出処理を示す図である。

【図3】フォーム学習時の処理の第1のフローチャートである。

【図4】運用時の処理の第1のフローチャートである。

【図5】情報処理装置の構成図である。

【図6】フォーム学習時の処理の第2のフローチャートである。

【図7】罫線構造抽出処理を示す図である。

【図8】管理情報位置指定処理を示す図である。

【図9】第1の大分類用罫線特徴を示す図である。

【図10】第2の大分類用罫線特徴を示す図である。

【図11】第3の大分類用罫線特徴を示す図である。

【図12】第4の大分類用罫線特徴を示す図である。

【図13】交点列の抽出方法を示す図である。

【図14】交点列を示す図である。

【図15】CROSS RATIO 算出処理のフローチャートである。

【図16】CROSS RATIO を用いた外郭罫線特徴を示す図である。

【図17】運用時の処理の第2のフローチャートである。

【図18】DPマッチングを示す図である。

【図19】DPマッチング処理のフローチャートである。

【図20】管理情報位置算出処理のフローチャート（その1）である。

【図21】管理情報位置算出処理のフローチャート（その2）である。

【図22】管理情報位置算出処理のフローチャート（その3）である。

【図23】ユーザ登録モードと自動学習モードを備えた

管理情報抽出処理を示す図である。

【図24】表内管理情報抽出処理のフローチャートである。

【図25】罫線のない文書画像の管理情報抽出処理のフローチャートである。

【図26】管理情報格納処理のフローチャートである。

【図27】管理情報格納テーブルを示す図である。

【図28】管理情報検索処理のフローチャートである。

【図29】連合グラフを示す図である。

【図30】フォーム識別処理のフローチャートである。

【図31】基準幅、基準高、基準点を示す図である。

【図32】横罫線を示す図である。

【図33】縦罫線を示す図である。

【図34】横罫線の詳細情報を示す図である。

【図35】縦罫線の詳細情報を示す図である。

【図36】モデルマッチング処理のフローチャートである。

【図37】マッチングテーブルを示す図である。

【図38】しきい値の関数を示す図である。

【図39】順序関係が逆になる場合を示す図である。

【図40】対応罫線が重複する場合を示す図である。

【図41】最適パス集合が示す罫線の対応関係を示す図である。

【図42】ノード配置処理のフローチャートである。

【図43】パス生成処理のフローチャート（その1）である。

【図44】パス生成処理のフローチャート（その2）である。

【図45】記憶部のノード列を示す図である。

【図46】詳細情報を用いた判定処理を示す図である。

【図47】最適パス集合決定処理のフローチャートである。

【図48】ノード数更新処理のフローチャートである。

【図49】先願のタイトル抽出装置の構成図である。

【図50】表形式文書を示す図である。

【符号の説明】

- 1 文字領域生成部
- 2 文字列領域生成部
- 3 タイトル抽出部
- 40 11 承認欄
- 12 項目名
- 13 タイトル
- 21 ユーザ登録手段
- 22 計算手段
- 23 辞書手段
- 24 照合手段
- 25 抽出手段
- 26 格納手段
- 27 グループ生成手段
- 28 検索手段

31 レイアウト辞書
 32、71、72 未知文書
 33 文字列領域
 41 CPU
 42 メモリ
 43 入力装置
 44 出力装置
 45 外部記憶装置
 46 媒体駆動装置
 47 ネットワーク接続装置
 48 光电変換装置
 49 バス
 50 可搬記録媒体

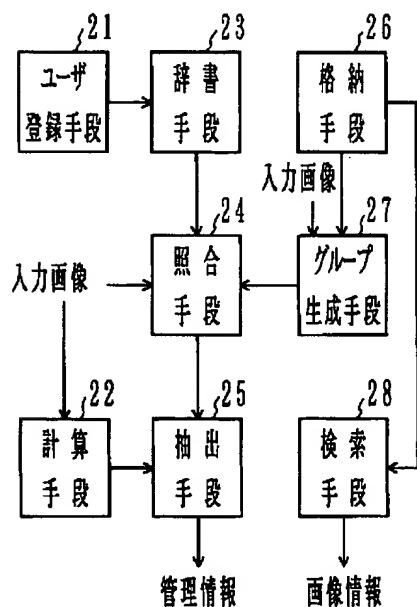
50' データベース
 51 矩形セル
 52、53、54、55、56、57、58、59、8
 5 頂点
 61、62、63、64、65、66、67、68、6
 9、70 交点
 73 ユーザ辞書
 74、75、76、77 処理結果
 81、82、83、84、86、87 罫線
 10 91、92、93、94、95、96、101、10
 2、103、104 罫線矩形
 97、98、99 ノード

【図1】

【図2】

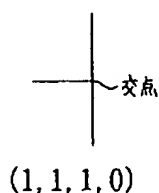
【図14】

本発明の原理図

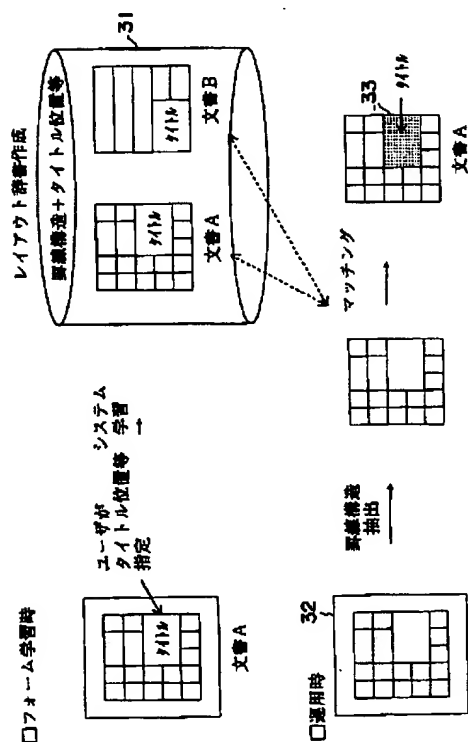


【図10】

第2の大分類用罫線特徴を示す図



管理情報抽出処理を示す図



交点列を示す図

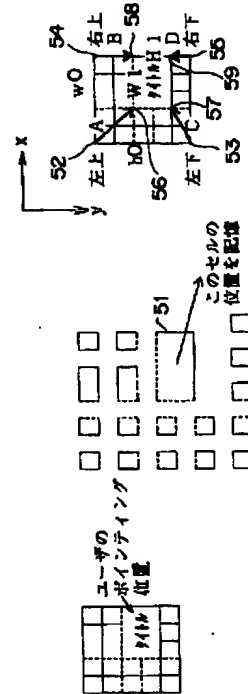
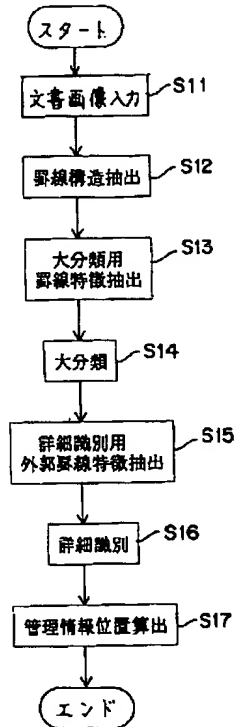
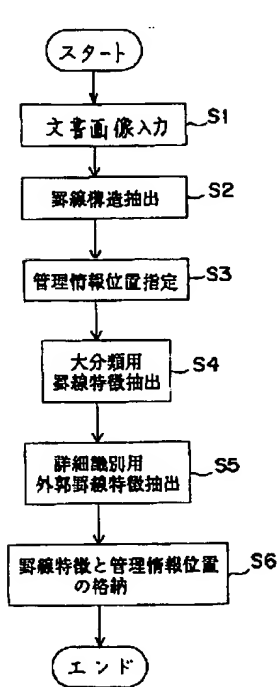


【図3】

【図4】

【図8】

フォーム学習時の処理の第1のフローチャート 運用時の処理の第1のフローチャート 管理情報位置指定処理を示す図

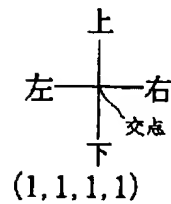
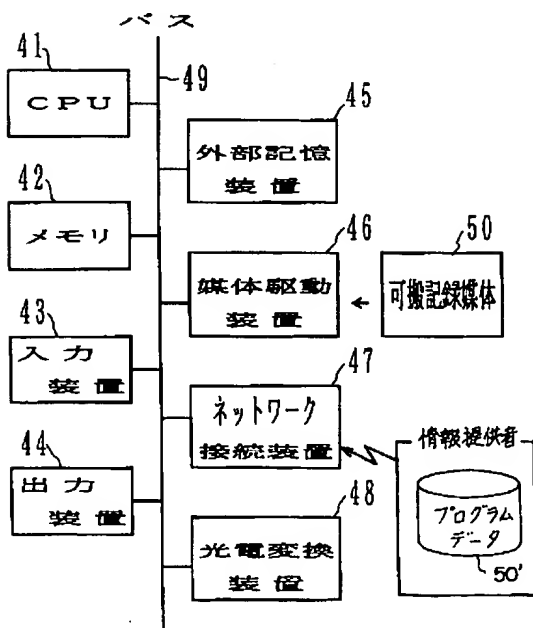


【図5】

【図9】

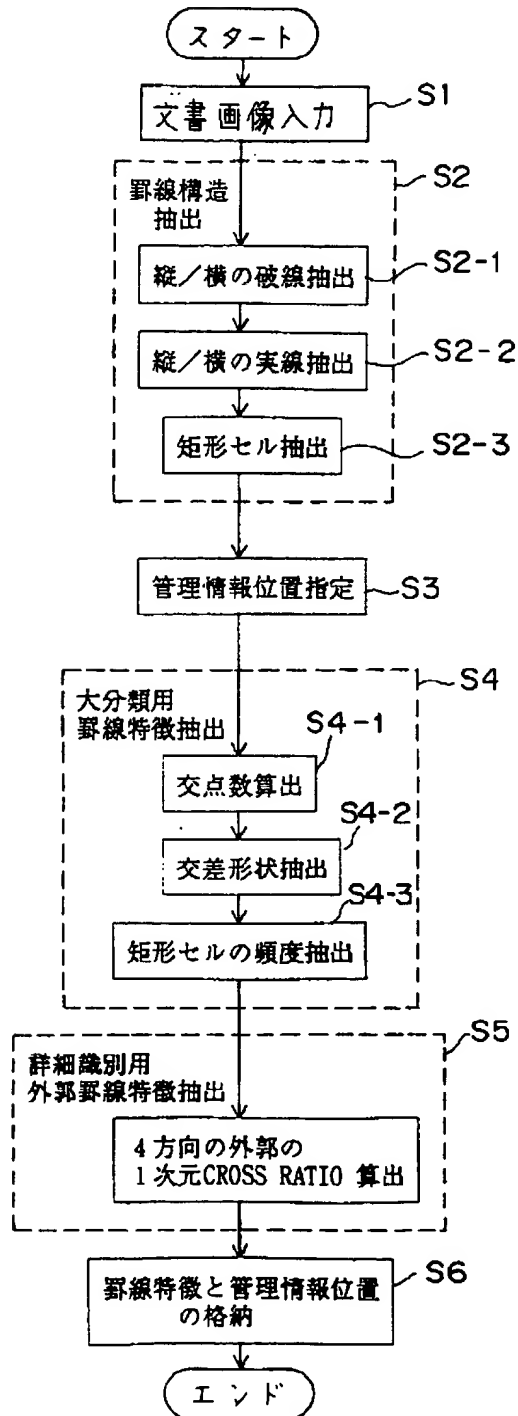
情報処理装置の構成図

第1の大分類用罫線特徴を示す図



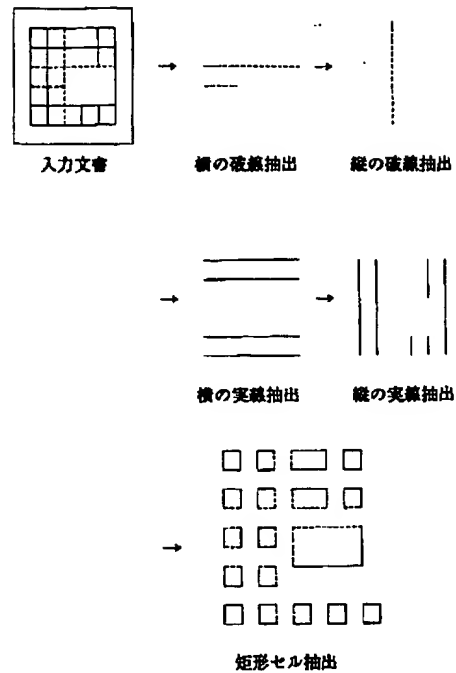
【図6】

フォーム学習時の処理の第2のフローチャート



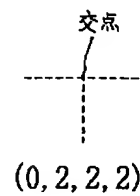
【図7】

罫線構造抽出処理を示す図



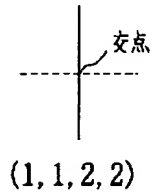
【図11】

第3の大分類用罫線特徴を示す図



【図12】

第4の大分類用算線特徴を示す図



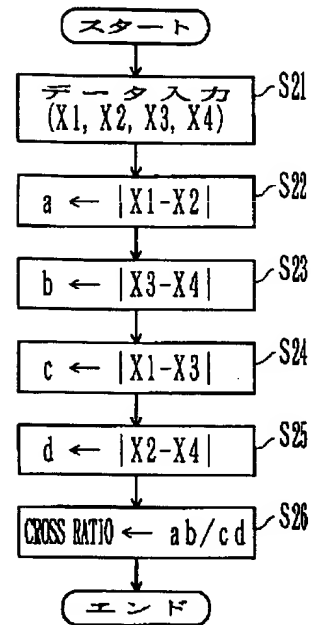
【図13】

交点列の抽出方法を示す図

	1 列目	2 列目	3 列目	4 列目	5 列目	6 列目
1 行目			66			
2 行目	61	62	63		64	65
3 行目			67			
4 行目			68			
5 行目			69			
6 行目			70			

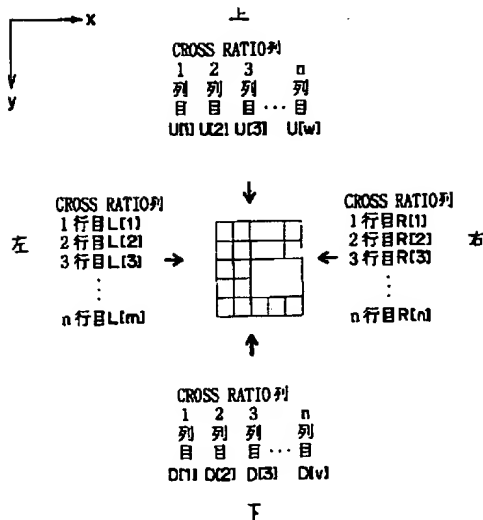
【図15】

CROSS RATIO 算出処理のフローチャート



【図16】

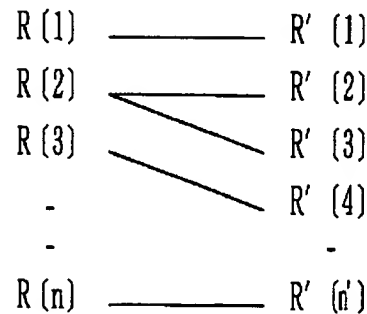
CROSS RATIO を用いた外郭算線特徴を示す図



【図18】

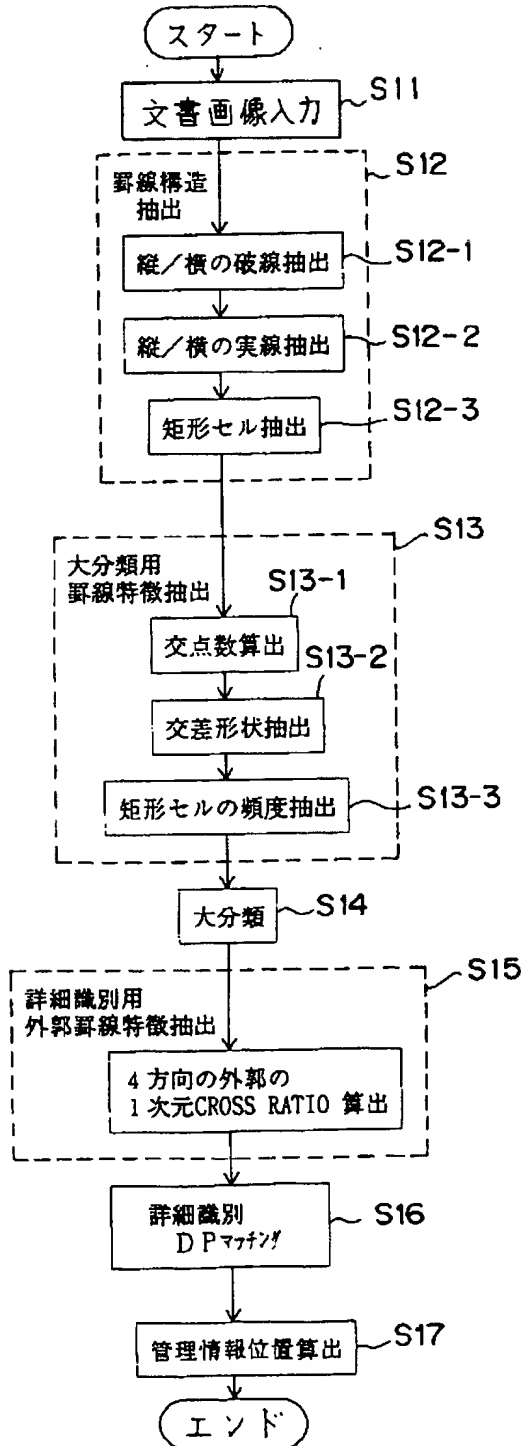
DP マッチングを示す図

入力フォーム 学習済みフォーム

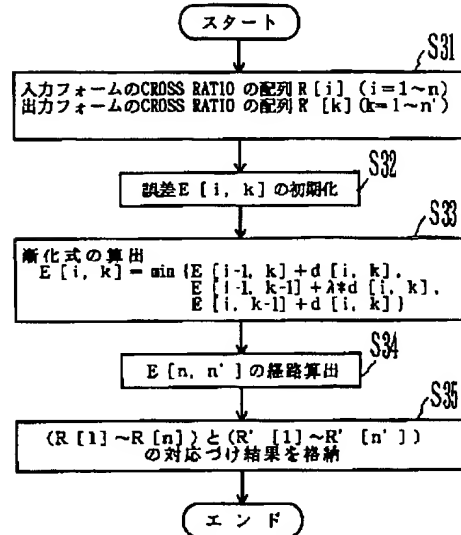


【図17】

運用時の処理の第2のフローチャート DPマッチング処理のフローチャート

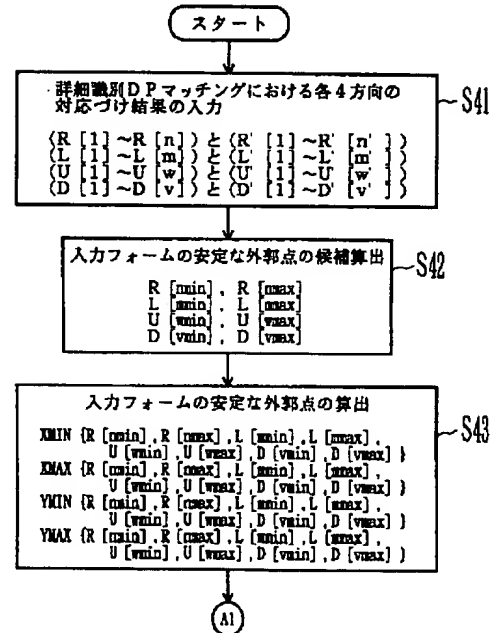


【図19】



【図20】

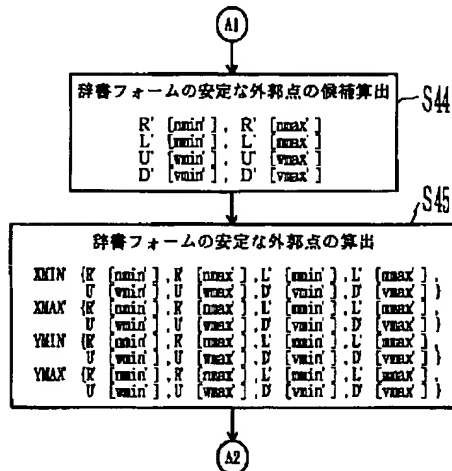
管理情報位置算出処理のフローチャート（その1）



【図 21】

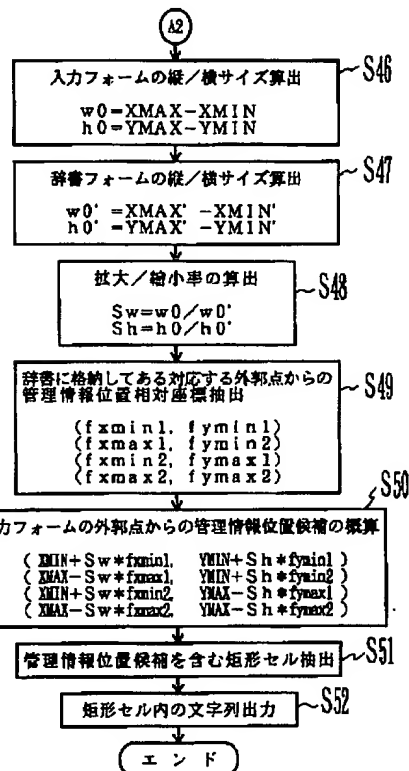
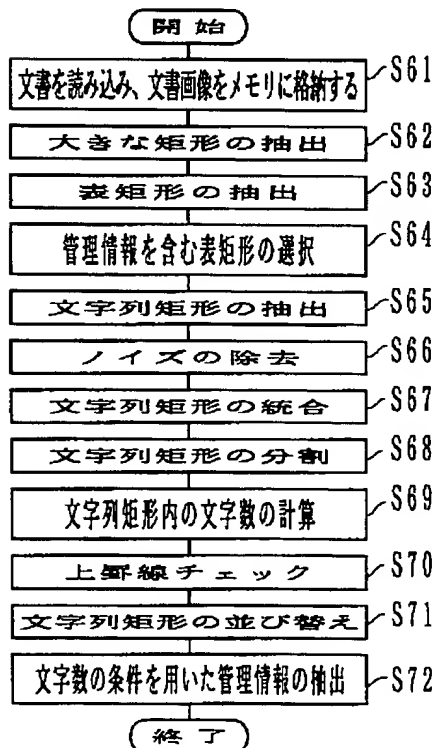
【図 22】

管理情報位置算出処理のフローチャート(その2) 管理情報位置算出処理のフローチャート(その3)



【図 24】

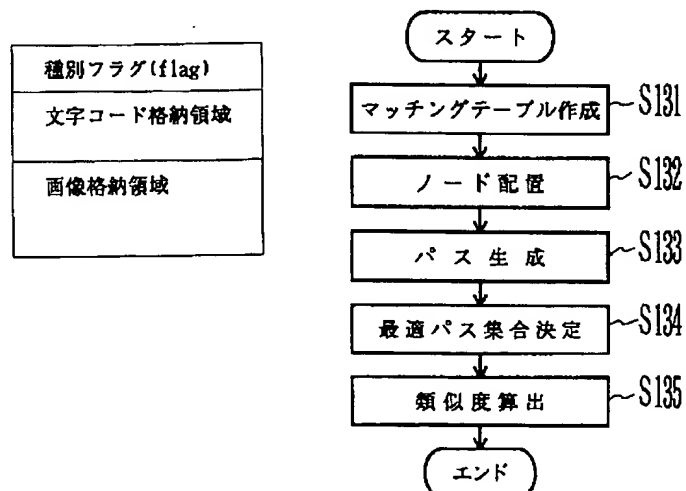
表内管理情報抽出処理のフローチャート



【図 27】

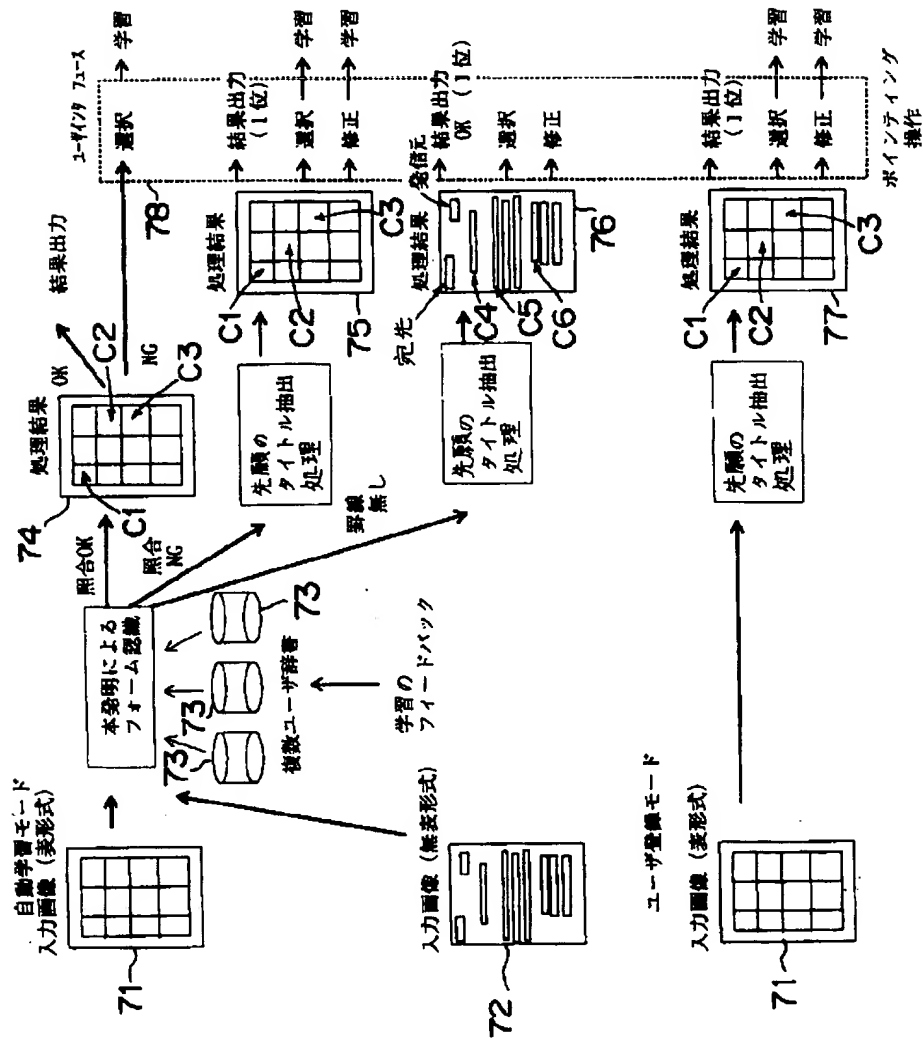
【図 36】

管理情報格納テーブルを示す図 モデルマッチング処理のフローチャート



【図23】

ユーザ登録モードと自動学習モードを備えた
管理情報抽出処理を示す図



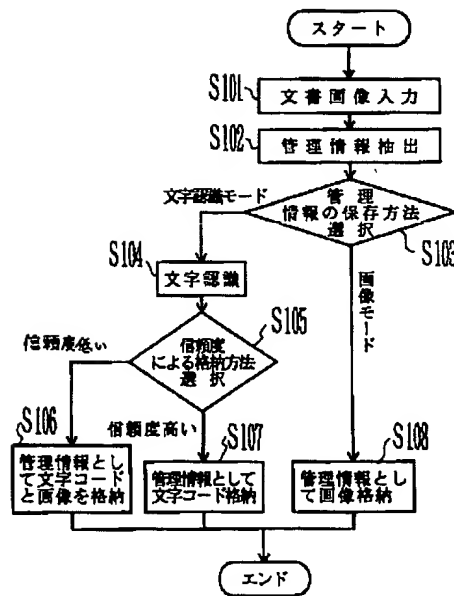
【図 25】

罫線のない文書画像の管理情報抽出処理のフローチャート



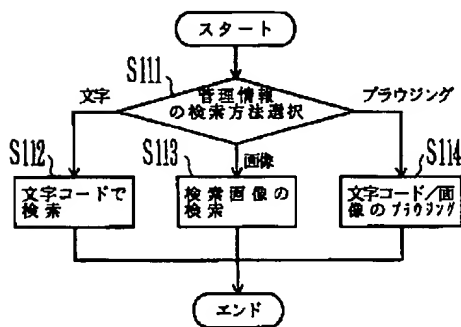
【図 26】

管理情報格納処理のフローチャート



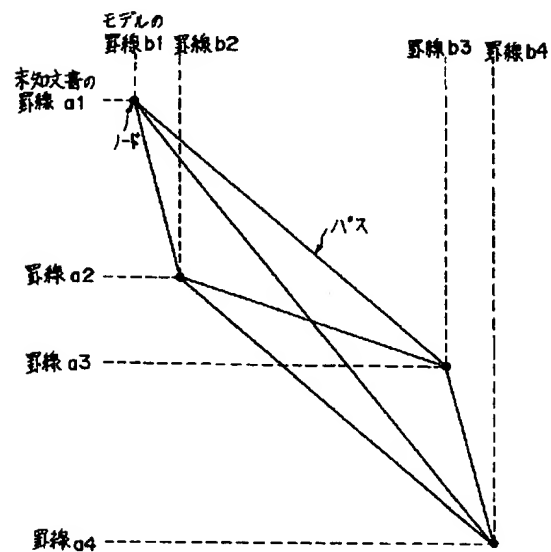
【図 28】

管理情報検索処理のフローチャート



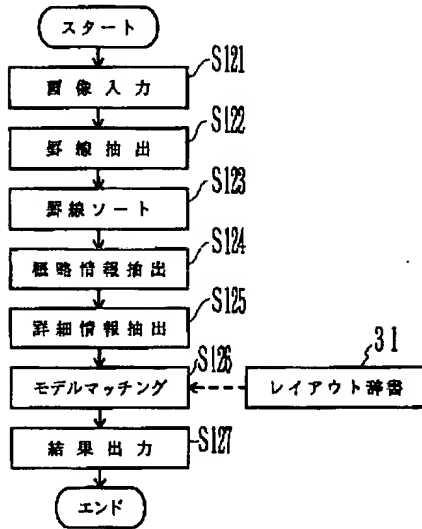
【図 29】

連合グラフを示す図



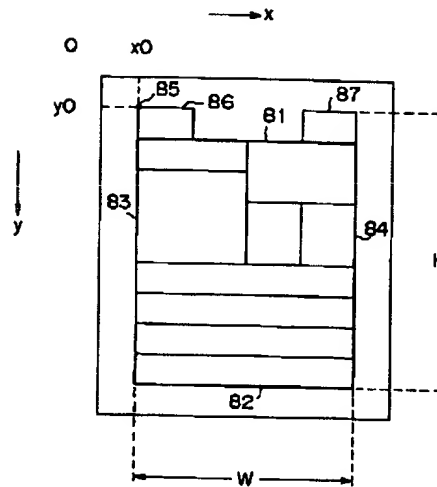
【図30】

フォーム識別処理のフローチャート



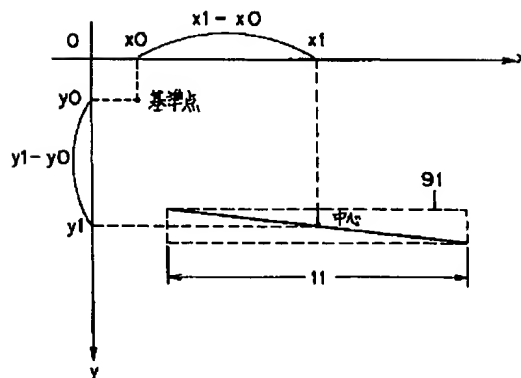
【図31】

基準幅、基準高、基準点を示す図



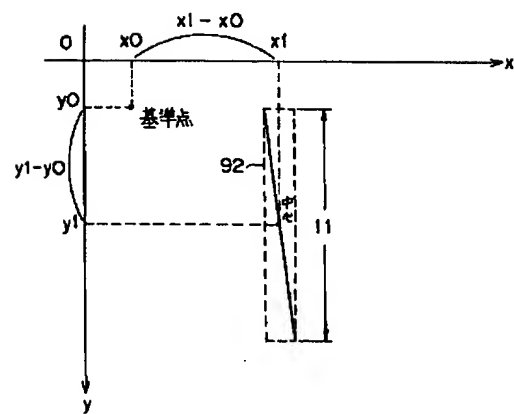
【図32】

横罫線を示す図



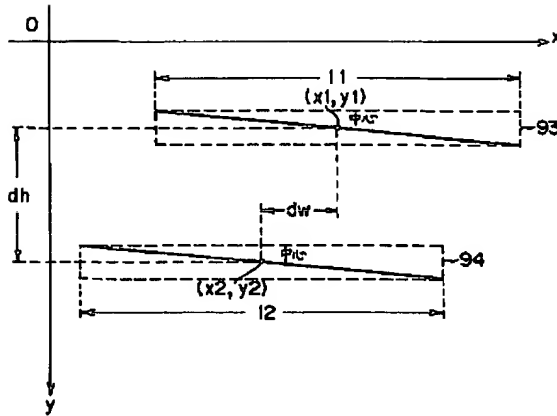
【図33】

縦罫線を示す図



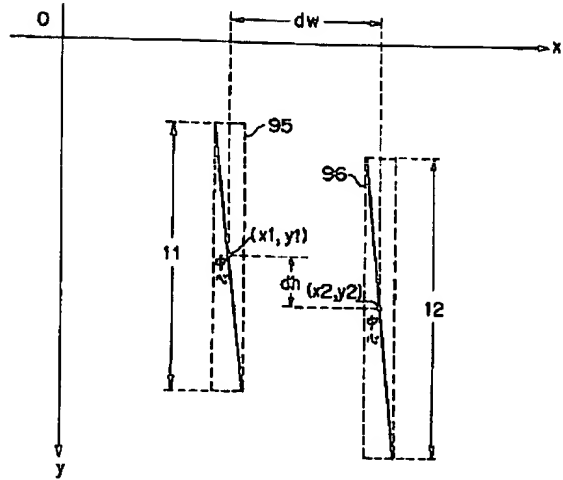
【図34】

横罫線の詳細情報を示す図



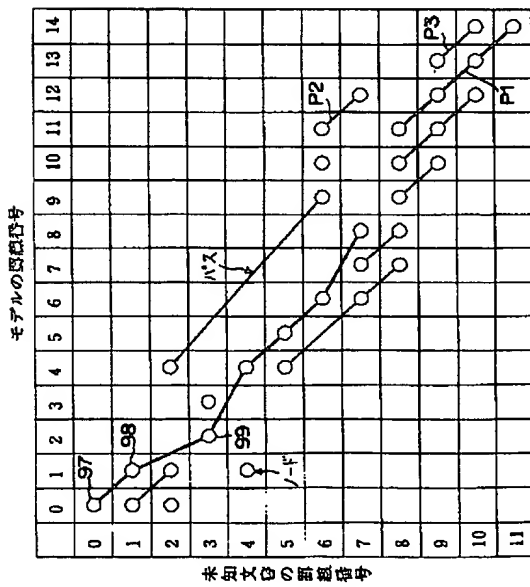
【図35】

縦罫線の詳細情報を示す図



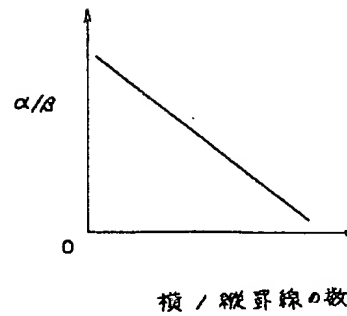
【図37】

マッチングテーブルを示す図



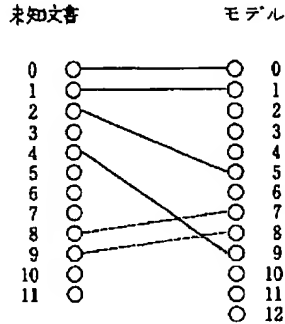
【図38】

しきい値の関数を示す図



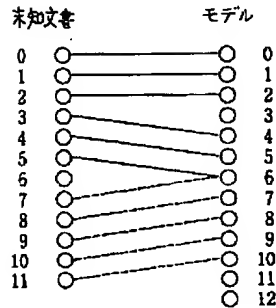
【図39】

順序関係が逆になる場合を示す図

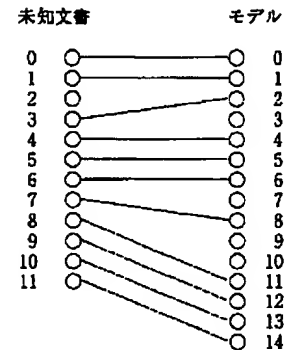


【図40】

対応異線が重複する場合を示す図

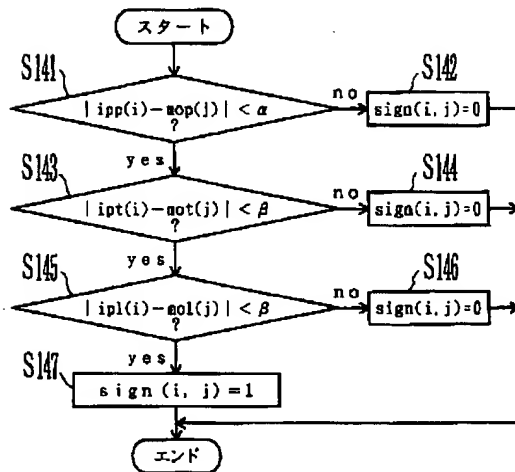


【図41】

最適パス集合が示す
異線の対応関係を示す図

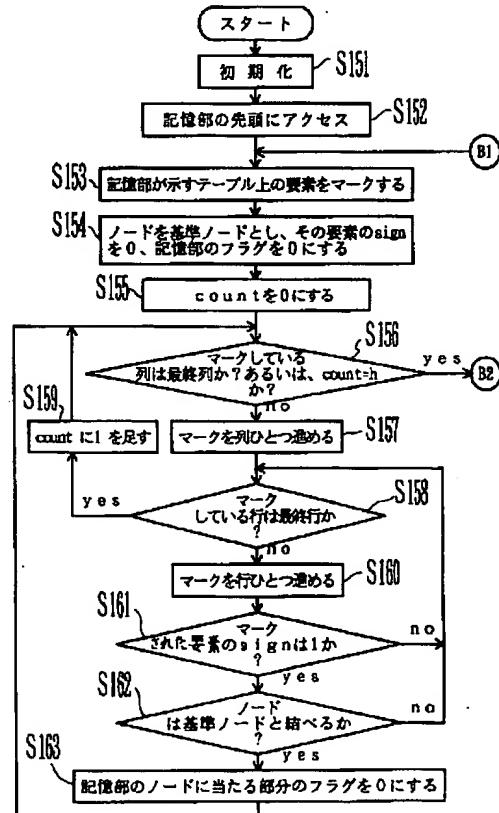
【図42】

ノード配置処理のフローチャート



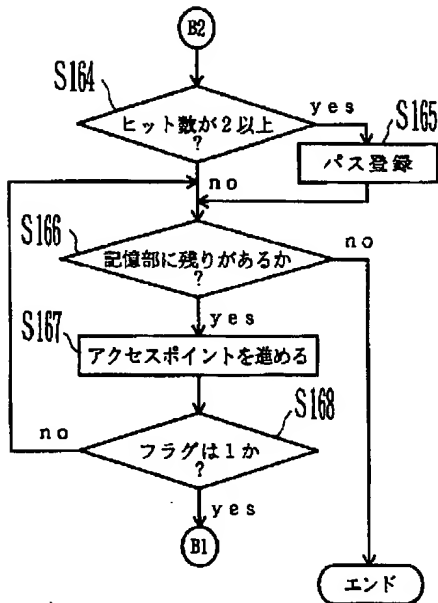
【図43】

パス生成処理のフローチャート (その1)



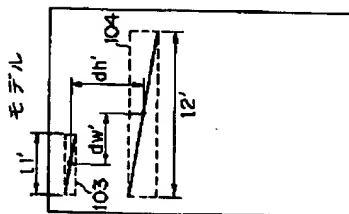
【図 4 4】

パス生成処理のフローチャート (その2)

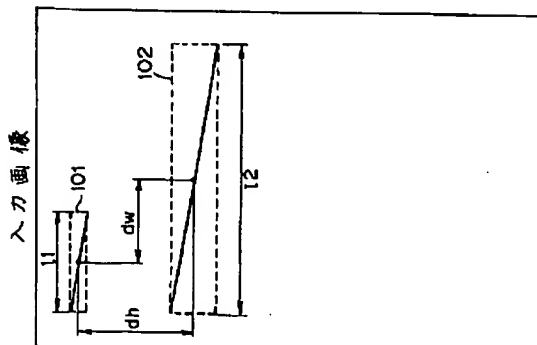


【図 4 6】

詳細情報を用いた判定処理を示す図



相似



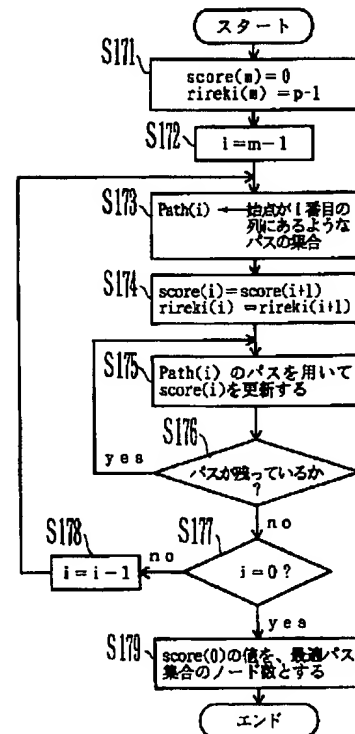
【図 4 5】

記憶部のノード列を示す図

ノードの位置	フラグ
< 0, 0 >	1
< 1, 0 >	1
< 1, 1 >	1
< 2, 0 >	1
⋮	⋮
< 11, 14 >	1

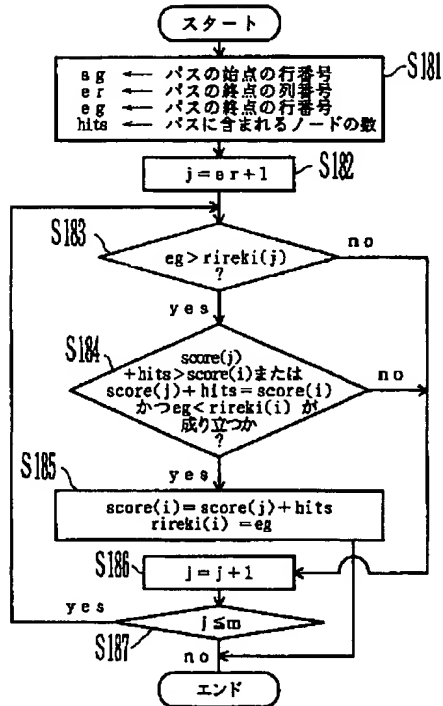
【図 4 7】

最適パス集合決定処理のフローチャート



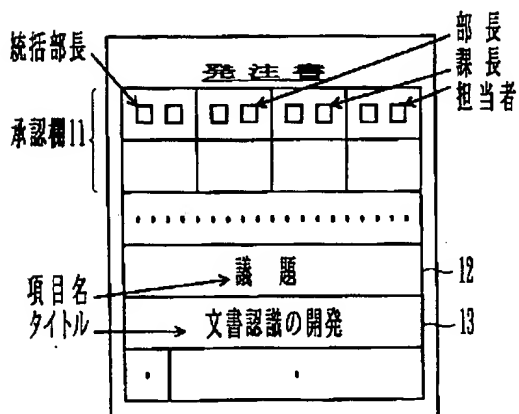
【図48】

ノード数更新処理のフローチャート



【図50】

表形式文書を示す図



【図49】

先願のタイトル抽出装置の構成図

